

OM protein - protein search, using sw model

Run on: September 17, 2004, 16:42:05 ; Search time 80.461 Seconds  
 (without alignments)  
 484.602 Million cell updates/sec

Title: US-10-057-890A-10  
 Perfect score: 797  
 Sequence: 1 MDYQVSSPIYDINYYTSEPC.....GLNNCSSSNRLDGHQRVHAA 138

Scoring table: BLOSUM62  
 Gapop 10.0 , Gapext 0.5

Searched: 1586107 seqs, 282547505 residues

Total number of hits satisfying chosen parameters: 1586107

Minimum DB seq length: 0  
 Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
 Maximum Match 100%  
 Listing first 45 summaries

Database : A\_Geneseq\_29Jan04:\*  
 1: geneseqp1980s:\*  
 2: geneseqp1990s:\*  
 3: geneseqp2000s:\*  
 4: geneseqp2001s:\*  
 5: geneseqp2002s:\*  
 6: geneseqp2003as:\*  
 7: geneseqp2003bs:\*  
 8: geneseqp2004s:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	% Query		DB	ID	Description
		Match	Length			
1	797	100.0	138	5	ABG32539	Abg32539 Human CCR
2	797	100.0	157	5	ABG32540	Abg32540 Human CCR
3	379	47.6	352	2	AAW27407	Aaw27407 Human CCR
4	379	47.6	352	2	AAW27123	Aaw27123 Human che
5	379	47.6	352	2	AAW23835	Aaw23835 Human CC
6	379	47.6	352	2	AAW88232	Aaw88232 HIV-1 co-
7	379	47.6	352	4	AAE07048	Aae07048 Human G-p
8	379	47.6	352	4	AAG80111	Aag80111 Human CCR
9	379	47.6	352	4	AAE04321	Aae04321 Human che

10	379	47.6	352	4	AAE07039	Aae07039 Human G-p
11	379	47.6	352	4	AAB46858	Aab46858 Human HDG
12	379	47.6	352	4	ABB56342	Abb56342 Non-endog
13	379	47.6	352	4	AAB83354	Aab83354 Human CCR
14	379	47.6	352	4	AAB82948	Aab82948 Human HIV
15	379	47.6	352	5	AAU97152	Aau97152 Human G-p
16	379	47.6	352	5	AAM52829	Aam52829 Human CCR
17	379	47.6	352	5	AAM52828	Aam52828 Human CC
18	379	47.6	352	5	ABG70597	Abg70597 Human G-p
19	379	47.6	352	5	ABG92883	Abg92883 Human imm
20	379	47.6	352	5	AAE25811	Aae25811 Human G-p
21	379	47.6	352	5	ABB81054	Abb81054 G-protein
22	379	47.6	352	5	ABB08343	Abb08343 Human che
23	379	47.6	352	6	ABG75540	Abg75540 Human G-p
24	379	47.6	352	6	ABR58602	Abr58602 Human can
25	379	47.6	352	6	AAO29514	Aao29514 Human C-C
26	379	47.6	352	6	ABU61654	Abu61654 Human G-p
27	379	47.6	352	6	ABP97728	Abp97728 Amino aci
28	379	47.6	352	6	ABP81933	Abp81933 Human C-C
29	379	47.6	352	7	ADC03341	Adc03341 Human che
30	379	47.6	439	2	AAy41280	Aay41280 Fusion pr
31	374	46.9	371	2	AAW23834	Aaw23834 Human CC
32	373	46.8	352	2	AAW07602	Aaw07602 Human G-p
33	373	46.8	352	3	AAy80128	Aay80128 Human G-p
34	373	46.8	352	4	AAE07046	Aae07046 Human G-p
35	373	46.8	352	4	AAE07037	Aae07037 Human G-p
36	373	46.8	352	5	AAU97150	Aau97150 Human G-p
37	373	46.8	352	5	ABG92880	Abg92880 Human G-p
38	373	46.8	352	5	AAE25808	Aae25808 Human G-p
39	363	45.5	352	2	AAW27125	Aaw27125 Macaque c
40	363	45.5	352	7	ADC03359	Adc03359 Macaque c
41	359	45.0	352	4	AAG79089	Aag79089 Amino aci
42	272	34.1	354	2	AAW54037	Aaw54037 Mouse CC-
43	258	32.4	184	2	AAW27406	Aaw27406 Inactive
44	258	32.4	215	2	AAW27408	Aaw27408 Inactive
45	258	32.4	215	2	AAW88238	Aaw88238 HIV-1 co-

# ALIGNMENTS

## RESULT 1

ABG32539

ID ABG32539 standard; protein; 138 AA.

XX

AC ABG32539;

XX

DT 15-NOV-2002 (first entry)

XX

DE Human CCR5-based scaffolded fusion protein #1.

XX

KW Scaffolded protein; CCR5; HIV; human immunodeficiency virus infection;

KW ECD; extracellular domain; metal chelating motif; zinc finger protein;

KW integral membrane protein; soluble loop; intracellular domain; ICD;

KW gene therapy; immunogen; viral infection; human.

XX

OS Homo sapiens.

OS Synthetic.  
 XX  
 PN WO200260477-A1.  
 XX  
 PD 08-AUG-2002.  
 XX  
 PF 29-JAN-2002; 2002WO-US002377.  
 XX  
 PR 31-JAN-2001; 2001US-0265782P.  
 PR 31-JAN-2001; 2001US-0265858P.  
 XX  
 PA (HUMA-) HUMAN GENOME SCI INC.  
 XX  
 PI Coleman TA, Mansfield B;  
 XX  
 DR WPI; 2002-643357/69.  
 XX  
 PT Novel scaffolded fusion polypeptide useful for therapeutic purposes or  
 PT for screening molecules that bind/activate/inhibit/modulate the  
 PT polypeptide, comprises a functional polypeptide domain fused to a  
 PT scaffold domain.  
 XX  
 PS Example 1; Page 21; 64pp; English.  
 XX  
 CC The invention relates to a scaffolded fusion polypeptide comprising a  
 CC functional polypeptide domain fused to a scaffold domain, where the  
 CC functional polypeptide domain corresponds to a soluble loop of an  
 CC integral membrane protein (e.g. human CCR5, a transmembrane receptor  
 CC involved in HIV (human immunodeficiency virus) infection). Also included  
 CC are; (1) a polypeptide comprising a scaffold domain; (2) a nucleic acid  
 CC encoding the fusion polypeptide; (3) a vector cassette for the expression  
 CC of the fusion polypeptide comprising an expression region operably linked  
 CC to a promoter, where the expression region comprises a number of  
 CC cassettes, each of which encodes a module, domain or strand of the fusion  
 CC polypeptide and (4) a host cell comprising the vector or nucleic acid.  
 CC The fusion polypeptide is useful for screening molecules that  
 CC bind/activate/inhibit/modulate the fusion polypeptide, by expressing the  
 CC fusion polypeptide from and identifying a molecule that binds to the  
 CC fusion polypeptide. The fusion polypeptide is useful in diagnostic  
 CC methods, in assays to identify compounds that interact with loops of  
 CC fragments of an extracellular domain (ECD) or an intracellular domain  
 CC (ICD) or to rapidly assay the function of mutated portions of mutant  
 CC integral membrane proteins without having to produce significant  
 CC quantities of the entire mutant integral membrane protein, to generate  
 CC antibodies that recognise the integral membrane proteins from which they  
 CC are designed, to competitively bind the ligand of a naturally occurring  
 CC receptor in vitro or in vivo, to display and/or screen soluble domains  
 CC from protein such as integral membrane proteins, to probe the structure  
 CC of ECD or ICD, or both, of an integral protein membrane, to modulate the  
 CC activity of a receptor in vivo, and for treating or preventing viral  
 CC infection, preferably human HIV infection e.g. by gene therapy using the  
 CC encoding nucleic acid. The present sequence is a scaffolded protein based  
 CC on the ECD region of human CCR5 (not defined)  
 XX  
 SQ Sequence 138 AA;

Query Match

100.0%; Score 797; DB 5; Length 138;

Best Local Similarity 100.0%; Pred. No. 1.8e-75;  
Matches 138; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
QY      1 MDYQVSSPIYDINYYTSEPCQKINVKQIAAYKCGLCAAQWDFGNTMCQHQRVHGHHS 60
      ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1 MDYQVSSPIYDINYYTSEPCQKINVKQIAAYKCGLCAAQWDFGNTMCQHQRVHGHHS 60

QY     61 YKCGLTRSQKEGLHYTCSSHPYSQYQFWKNFQTLKIHQRVHGGGGSYKCGLCQEFFGL 120
      ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     61 YKCGLTRSQKEGLHYTCSSHPYSQYQFWKNFQTLKIHQRVHGGGGSYKCGLCQEFFGL 120

QY    121 NNCSSSNRLDGHQRVHAA 138
      ||||||||||||||||
Db    121 NNCSSSNRLDGHQRVHAA 138
```

RESULT 2

ABG32540

ID ABG32540 standard; protein; 157 AA.

XX

AC ABG32540;

XX

DT 15-NOV-2002 (first entry)

XX

DE Human CCR5-based scaffolded fusion protein #2.

XX

KW Scaffolded protein; CCR5; HIV; human immunodeficiency virus infection;  
KW ECD; extracellular domain; metal chelating motif; zinc finger protein;  
KW integral membrane protein; soluble loop; intracellular domain; ICD;  
KW gene therapy; immunogen; viral infection; human.

XX

OS Homo sapiens.

OS Synthetic.

XX

FH Key Location/Qualifiers

FT Peptide 1. .19

FT /label= Signal\_peptide

FT Protein 20. .157

FT /label= Mature\_scaffolded\_protein

XX

PN WO200260477-A1.

XX

PD 08-AUG-2002.

XX

PF 29-JAN-2002; 2002WO-US002377.

XX

PR 31-JAN-2001; 2001US-0265782P.

PR 31-JAN-2001; 2001US-0265858P.

XX

PA (HUMA-) HUMAN GENOME SCI INC.

XX

PI Coleman TA, Mansfield B;

XX

DR WPI; 2002-643357/69.

XX

PT Novel scaffolded fusion polypeptide useful for therapeutic purposes or  
PT for screening molecules that bind/activate/inhibit/modulate the



PT polypeptide, comprises a functional polypeptide domain fused to a  
PT scaffold domain.  
XX  
PS Example 2; Page 41; 64pp; English.  
XX  
CC The invention relates to a scaffolded fusion polypeptide comprising a  
CC functional polypeptide domain fused to a scaffold domain, where the  
CC functional polypeptide domain corresponds to a soluble loop of an  
CC integral membrane protein (e.g. human CCR5, a transmembrane receptor  
CC involved in HIV (human immunodeficiency virus) infection). Also included  
CC are; (1) a polypeptide comprising a scaffold domain; (2) a nucleic acid  
CC encoding the fusion polypeptide; (3) a vector cassette for the expression  
CC of the fusion polypeptide comprising an expression region operably linked  
CC to a promoter, where the expression region comprises a number of  
CC cassettes, each of which encodes a module, domain or strand of the fusion  
CC polypeptide and (4) a host cell comprising the vector or nucleic acid.  
CC The fusion polypeptide is useful for screening molecules that  
CC bind/activate/inhibit/modulate the fusion polypeptide, by expressing the  
CC fusion polypeptide from and identifying a molecule that binds to the  
CC fusion polypeptide. The fusion polypeptide is useful in diagnostic  
CC methods, in assays to identify compounds that interact with loops of  
CC fragments of an extracellular domain (ECD) or an intracellular domain  
CC (ICD) or to rapidly assay the function of mutated portions of mutant  
CC integral membrane proteins without having to produce significant  
CC quantities of the entire mutant integral membrane protein, to generate  
CC antibodies that recognise the integral membrane proteins from which they  
CC are designed, to competitively bind the ligand of a naturally occurring  
CC receptor in vitro or in vivo, to display and/or screen soluble domains  
CC from protein such as integral membrane proteins, to probe the structure  
CC of ECD or ICD, or both, of an integral protein membrane, to modulate the  
CC activity of a receptor in vivo, and for treating or preventing viral  
CC infection, preferably human HIV infection e.g. by gene therapy using the  
CC encoding nucleic acid. The present sequence is a scaffolded protein based  
CC on the ECD region of human CCR5 (not defined)  
XX  
SQ Sequence 157 AA;

Query Match 100.0%; Score 797; DB 5; Length 157;  
Best Local Similarity 100.0%; Pred. No. 2.1e-75;  
Matches 138; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MDYQVSSPIYDINYYTSEPCQKINVKQIAAYKCGLCAAQWDFGNTMCQHQRVHGHHS 60  
|  
Db 20 MDYQVSSPIYDINYYTSEPCQKINVKQIAAYKCGLCAAQWDFGNTMCQHQRVHGHHS 79  
QY 61 YKCGLCTRSQKEGLHYTCSSHPYSQYQFWKNFQTLKIHQRVHGGGGSYKCGLCQEFFGL 120  
|  
Db 80 YKCGLCTRSQKEGLHYTCSSHPYSQYQFWKNFQTLKIHQRVHGGGGSYKCGLCQEFFGL 139  
QY 121 NNCSSSNRLDGHQRVHAA 138  
|  
Db 140 NNCSSSNRLDGHQRVHAA 157

RESULT 3  
AAW27407  
ID AAW27407 standard; protein; 352 AA.

XX  
 AC AAW27407;  
 XX  
 DT 14-APR-1998 (first entry)  
 XX  
 DE Human CCR5.  
 XX  
 KW Human Cys-Cys chemokine receptor 5; CCR5; human immunodeficiency virus;  
 KW type 1; type 2; HIV-1; HIV-2; diagnosis; treatment; prevention;  
 KW inflammatory disease; rheumatoid arthritis; glomerulonephritis; asthma;  
 KW idiopathic pulmonary fibrosis; psoriasis; viral infection; cancer;  
 KW atherosclerosis; autoimmune disorder.  
 XX  
 OS Homo sapiens.  
 XX  
 PN WO9732019-A2.  
 XX  
 PD 04-SEP-1997.  
 XX  
 PF 28-FEB-1997; 97WO-BE000023.  
 XX  
 PR 01-MAR-1996; 96EP-00870021.  
 PR 06-AUG-1996; 96EP-00870102.  
 XX  
 PA (EURO-) EUROSCREEN SA.  
 XX  
 PI Samson M, Parmentier M, Vassart G, Libert F;  
 XX  
 DR WPI; 1997-479829/44.  
 DR N-PSDB; AAT90117.  
 XX  
 PT Active and inactive forms of human CC chemokine receptor CCR-5 - useful  
 PT to diagnose, prevent and/or treat inflammatory disorders, autoimmune  
 PT disease and viral infection.  
 XX  
 PS Claim 4; Fig 1b-c; 94pp; English.  
 XX  
 CC The present sequence is human CC (Cys-Cys) chemokine receptor 5 (CCR5),  
 CC which is stimulated by MIP-1 alpha, MIP-1 beta or RANTES chemokines, but  
 CC not by monocyte chemoattractant protein 1 (MCP-1), MCP-2, MCP-3,  
 CC interleukin-8 (IL-8) or growth related gene product alpha (GRO alpha)  
 CC chemokines. Active CCR-5 is also a receptor of human immunodeficiency  
 CC virus type 1 or type 2 (HIV-1 or HIV-2). CCR5 or its cDNA can used to  
 CC diagnose, treat and/or prevent inflammatory diseases, e.g. rheumatoid  
 CC arthritis, glomerulonephritis, asthma, idiopathic pulmonary fibrosis and  
 CC psoriasis, viral infections, especially HIV-1 or HIV-2 infection, cancer,  
 CC atherosclerosis and autoimmune disorders  
 XX  
 SQ Sequence 352 AA;

Query Match 47.6%; Score 379; DB 2; Length 352;  
 Best Local Similarity 35.0%; Pred. No. 3.2e-31;  
 Matches 99; Conservative 7; Mismatches 27; Indels 150; Gaps 7;

Qy 1 MDYQVSSPIYDINYTTSEPCQKINVKQIAA----- 30  
 |||||  
 Db 1 MDYQVSSPIYDINYTTSEPCQKINVKQIAARLLPPLYSLVFIFGFVGNMLVILILINCKR 60



PF 20-DEC-1996; 96WO-US020759.  
 XX  
 PR 20-DEC-1995; 95US-00575967.  
 PR 07-JUN-1996; 96US-00661393.  
 XX  
 PA (ICOS-) ICOS CORP.  
 XX  
 PI Gray PW, Schweickart VL, Raport CJ;  
 XX  
 DR WPI; 1997-341689/31.  
 DR N-PSDB; AAT85161.  
 XX  
 PT New nucleic acid encoding chemokine receptors 88-2B and 88C - used to  
 PT modulate leukocyte trafficking, e.g. for treatment of inflammation,  
 PT tumours, viral infections, auto-immune diseases, etc.  
 XX  
 PS Claim 16; Page 47-48; 65pp; English.  
 XX  
 CC This polypeptide sequence comprises novel human chemokine receptor 88C, a  
 CC G protein coupled receptor that is involved in leukocyte trafficking. Its  
 CC amino sequence was deduced from a cDNA clone (AAT85161) isolated from a  
 CC macrophage library. It shows 62% identity to CCKR1. Chemokine receptor  
 CC 88-2B (see AAW27124) has also been identified. 88C and 88-2B receptors  
 CC and their polypeptide fragments can be produced in transformed host  
 CC cells. The receptors, peptides comprising one or more of the  
 CC extracellular or intracellular domains, and anti-receptor antibodies can  
 CC be used to modulate receptor activities, particularly ligand and G  
 CC protein binding, and are potentially potentially useful in the treatment  
 CC of atherosclerosis, rheumatoid arthritis, tumours, asthma, viral  
 CC infection, AIDS, inflammatory conditions, pathological immune response,  
 CC abnormal haematopoietic processes etc  
 XX  
 SQ Sequence 352 AA;

Query Match 47.6%; Score 379; DB 2; Length 352;  
 Best Local Similarity 35.0%; Pred. No. 3.2e-31;  
 Matches 99; Conservative 7; Mismatches 27; Indels 150; Gaps 7;

QY	1 MDYQVSSPIYDINYYTSEPCQKINVKQIAA-----	30
Db	1 MDYQVSSPIYDINYYTSEPCQKINVKQIAARLLPLYSLVFIFGFVGNMLVILILINCKR	60
QY	31 -----YKCGLC-----AAAQWDFGNTMCQHQRVHGHHSYKCG---	64
	:   :	
Db	61 LKSMTDIYLLNLAISDLFFLLTVPFWAHYAAAQWDFGNTMC--QLLTGLYFIGFFSGIFF	118
QY	65 -----LCTRSQKEGLHYTC	78
	:	
Db	119 IILLTIDRYLAVVHAVFALKARTVTEGVVTSVITWVAVFASLPGLIFTRSQKEGLHYTC	178
QY	79 SSHFPYSQYQFWKNFQTLKI-----HQRVHGG-----	105
	: :	
Db	179 SSHFPYSQYQFWKNFQTLKIVILGLVLPPLVMVICYSGILKTLLRCRNEKKRRAVRLIF	238
QY	106 -----GGSYKCGLC-----QEFFGLNNCSSSNRLDGHQRV	135
	:	
Db	239 TIMIVYFLFWAPYNIVLLNTFQEFFGLNNCSSSNRLDQAMQV	281

RESULT 5  
 AAW23835  
 ID AAW23835 standard; protein; 352 AA.  
 XX  
 AC AAW23835;  
 XX  
 DT 08-JUN-1998 (first entry)  
 XX  
 DE Human CC chemokine receptor 5 (CCR5).  
 XX  
 KW CC chemokine receptor 5; CCR5; G-protein coupled receptor;  
 KW human immunodeficiency virus; HIV; CD4; AIDS; therapy; transgenic animal.  
 XX  
 OS Homo sapiens.  
 XX  
 FH Key Location/Qualifiers  
 FT Domain 29. .55  
 FT /label= I  
 FT /note= "transmembrane domain"  
 FT Domain 104. .126  
 FT /label= III  
 FT /note= "transmembrane domain"  
 FT Region 109. .120  
 FT /note= "extracellular loop-1 (Claim 19)"  
 FT Domain 143. .171  
 FT /label= IV  
 FT /note= "transmembrane domain"  
 FT Region 187. .210  
 FT /note= "extracellular loop-2 (Claim 19)"  
 FT Domain 194. .219  
 FT /label= V  
 FT /note= "transmembrane domain"  
 FT Domain 238. .258  
 FT /label= VI  
 FT /note= "transmembrane domain"  
 FT Region 261. .276  
 FT /note= "extracellular loop-3 (Claim 19)"  
 FT Domain 277. .300  
 FT /label= VII  
 FT /note= "transmembrane domain"  
 XX  
 PN WO9745543-A2.  
 XX  
 PD 04-DEC-1997.  
 XX  
 PF 28-MAY-1997; 97WO-US009586.  
 XX  
 PR 28-MAY-1996; 96US-0018508P.  
 XX  
 PA (USSH ) US DEPT HEALTH & HUMAN SERVICES.  
 XX  
 PI Combadiere C, Feng Y, Berger EA, Alkhatib G, Murphy PM;  
 PI Broder CC, Kennedy PE;  
 XX  
 DR WPI; 1998-032650/03.

DR N-PSDB; AAT76920.  
 XX  
 PT CC chemokine receptor 5 polypeptide - used to inhibit membrane fusion  
 PT between HIV and a target cell.  
 XX  
 PS Claim 68; Fig 1C; 70pp; English.  
 XX  
 CC This protein sequence comprises of a novel human macrophage-selective CC  
 CC chemokine receptor that has been designated CCR5. The sequence was  
 CC deduced from an isolated cDNA clone (see AAT76920). An Ala127Leu variant  
 CC (see W238340 of CCR5 was also identified. The susceptibility of human  
 CC macrophages to HIV infection depends on cell surface expression of CD4  
 CC and CCR5. CCR5 is a member of the 7-transmembrane superfamily of G-  
 CC protein coupled cell surface molecules. It plays an essential role in the  
 CC membrane fusion step of infection by some HIV isolates. The establishment  
 CC of stable, non-human cell lines and transgenic mammals having cells that  
 CC coexpress human CD4 and CCR5 provides valuable tools for research of HIV  
 CC infection. Antibodies that bind to CCR5, CCR5 variants, and CCR5-binding  
 CC agents capable of blocking membrane fusion between HIV and target cells  
 CC represent potential anti-HIV therapeutics for macrophage tropic strains  
 CC of HIV  
 XX  
 SQ Sequence 352 AA;

Query Match 47.6%; Score 379; DB 2; Length 352;  
 Best Local Similarity 35.0%; Pred. No. 3.2e-31;  
 Matches 99; Conservative 7; Mismatches 27; Indels 150; Gaps 7;

```

Qy      1 MDYQVSSPIYDINYYTSEPCQKINVKQIAA----- 30
          ||||||||||||||||||||||||||||
Db      1 MDYQVSSPIYDINYYTSEPCQKINVKQIAARLLPPLYSLVFIFGFVGNMLVILILINCKR 60

Qy      31 -----YKCGLC-----AAAQWDFGNTMCQHQRVHGHHSYKCG--- 64
          |  |                               |||||  | : | : |
Db      61 LKSMTDIYLLNLAISDLFFLLTVPFWAHYAAAQWDFGNTMC--QLLTGLYFIGFFSGIFF 118

Qy      65 -----LCTRSQKEGLHYTC 78
          : |||||
Db      119 IILLTIDRYLAVVHAVFALKARTVTFGVVTSVITWVAVFASLPGIIFTRSQKEGLHYTC 178

Qy      79 SSHFPYSQYQFWKNFQTLKI-----HQRVHGG----- 105
          ||||||||||||||||
Db      179 SSHFPYSQYQFWKNFQTLKIVILGLVPLLMVICYSGILKTLLRCRNEKKRHRVRLIF 238

Qy      106 -----GGSYKCGLC-----QEFFGLNNCSSSNRLDGHQRV 135
          |  |                               |||||  :|
Db      239 TIMIVYFLFWAPYNIVLLLNLFQEFFGLNNCSSSNRLDQAMQV 281

```

# RESULT 6

AAW88232

ID AAW88232 standard; protein; 352 AA.

XX

AC AAW88232;

XX

DT 15-MAR-1999 (first entry)

XX

DE HIV-1 co-receptor CCR5.  
 XX  
 KW HIV-1; CCR5; CCR5m303; co-receptor; infection; diagnosis; AIDS;  
 KW gene therapy; human.  
 XX  
 OS Homo sapiens.  
 XX  
 FH Key Location/Qualifiers  
 FT Domain 32. .56  
 FT /note= "transmembrane domain 1"  
 FT Domain 67. .87  
 FT /note= "transmembrane domain 2"  
 FT Misc-difference 101  
 FT /note= "corresponds to TGT (Cys) in wild-type CCR5, TGA  
 FT (Stop) in CCR5m303"  
 FT Domain 103. .124  
 FT /note= "transmembrane domain 3"  
 FT Domain 142. .167  
 FT /note= "transmembrane domain 4"  
 FT Domain 200. .223  
 FT /note= "transmembrane domain 5"  
 FT Domain 236. .260  
 FT /note= "transmembrane domain 6"  
 FT Domain 275. .301  
 FT /note= "transmembrane domain 7"  
 XX  
 PN WO9854317-A1.  
 XX  
 PD 03-DEC-1998.  
 XX  
 PF 29-MAY-1998; 98WO-EP003437.  
 XX  
 PR 30-MAY-1997; 97US-0048057P.  
 XX  
 PA (MOND-) FOND MONDIALE RECH & PREVENTION SIDA.  
 XX  
 PI Beretta A, Quillent C, Arenzana Siesdedos F, Braun J;  
 XX  
 DR WPI; 1999-059835/05.  
 DR N-PSDB; AAV84126.  
 XX  
 PT New CCR5 variant protein of the HIV-1 co-receptor - useful in developing  
 PT resistance of CCR5-expressing cells to HIV-1 infection.  
 XX  
 PS Disclosure; Page 34-35; 55pp; English.  
 XX  
 CC This is the amino acid sequence of wild-type human CCR5, which serves as  
 CC a co-receptor for infection by macrophage-tropic (M-tropic) strains of  
 CC HIV-1. The invention relates to the identification of a CCR5 variant (see  
 CC AAW88231), designated CCR5m303, comprising the first two transmembrane  
 CC domains of wild-type CCR5, but lacking transmembrane domains 3-7. The  
 CC presence of the CCR5m303 variant with the wild type CCR5 allele shows a  
 CC positive correlation with resistance to infection with M-tropic HIV-1  
 CC strains, and may indicate slower progression of the disease. The  
 CC detection of CCR5 variants may be used to identify individuals at lower  
 CC risk of infection relative to the general population who, if infected,  
 CC may exhibit slower progression to AIDS. Probes and primers (see AAV84127-

CC 36) are provided for use in diagnostic methods for detecting the presence  
 CC of such variants. A method is provided for inhibiting HIV-1 infection of  
 CC a cell expressing the CCR5 receptor. This involves introducing a nucleic  
 CC acid encoding a CCR5 variant into the cell, thereby reducing the number  
 CC of functional CCR5 molecules present on the cell surface  
 XX  
 SQ Sequence 352 AA;

Query Match 47.6%; Score 379; DB 2; Length 352;  
 Best Local Similarity 35.0%; Pred. No. 3.2e-31;  
 Matches 99; Conservative 7; Mismatches 27; Indels 150; Gaps 7;

```

QY      1 MDYQVSSPIYDINYYTSEPCQKINVKQIAA----- 30
          ||||||||||||||||||||
Db      1 MDYQVSSPIYDINYYTSEPCQKINVKQIAARLLPPLYSLVFIFGFVGNMLVILILINCKR 60

QY     31 -----YKGLC-----AAAQWDFGNTMCQHQRVHGHHSYKCG--- 64
          | | ||||||||| | : | : |
Db     61 LKSMTDIYLLNLAISDLFFLLTVPFWAHYAAAQWDFGNTMC--QLLTGLYFIGFFSGIFF 118

QY     65 -----LCTRSQKEGLHYTC 78
          : |||||||||
Db    119 IILLTIDRYLAVVHAVFALKARTVTFGVVTSVITWVAVFASLPGIIFTRSQKEGLHYTC 178

QY     79 SSHFPYSQYQFWKNEQTLKI-----HQRVHGG----- 105
          |||||||||||||||| : |
Db    179 SSHFPYSQYQFWKNEQTLKIVILGLVLPPLVMVICYSGILKTLLRCRNEKKRHRAVRLIF 238

QY    106 -----GGSYKGLC-----QEFFGLNNCSSSNRLDGHQRV 135
          | | ||||||||| : |
Db    239 TIMIVYFLFWAPYNIVLLLNTFQEFFGLNNCSSSNRLDQAMQV 281
  
```

# RESULT 7

AAE07048

ID AAE07048 standard; protein; 352 AA.

XX

AC AAE07048;

XX

DT 16-OCT-2001 (first entry)

XX

DE Human G-protein chemokine receptor (CCR5) HDGNR10 protein #2.

XX

KW Human; G-protein chemokine receptor; CCR5; HDGNR10; inflammation; HIV;  
 KW human immunodeficiency virus; antimicrobial; vasodilator; vulnerary;  
 KW cytostatic; immunosuppressive; nootropic; neuroprotective; gene therapy;  
 KW neurodegenerative disorder; Kaposi's sarcoma; autoimmune disease;  
 KW rheumatoid arthritis; cancer; breast; ovary; adrenal gland; bone marrow;  
 KW gastrointestinal tract; lung; liver; immune disorder; Addison's disease;  
 KW haemolytic anaemia; autoimmune thyroiditis; diabetes mellitus; allergy;  
 KW multiple sclerosis; ulcerative colitis; Crohn's disease; wound healing;  
 KW cardiovascular disorder; myocardial ischaemia.

XX

OS Homo sapiens.

XX

PN WO200158916-A2.

XX



PD 16-AUG-2001.  
 XX  
 PF 09-FEB-2001; 2001WO-US004153.  
 XX  
 PR 09-FEB-2000; 2000US-0181258P.  
 PR 09-MAR-2000; 2000US-0187999P.  
 PR 22-SEP-2000; 2000US-0234336P.  
 XX  
 PA (HUMA-) HUMAN GENOME SCI INC.  
 XX  
 PI Rosen CA, Roschke V, Li Y, Ruben SM;  
 XX  
 DR WPI; 2001-488966/53.  
 DR N-PSDB; AAD13299.  
 XX  
 PT Isolated nucleic acid encoding a human G-protein chemokine receptor  
 PT (CCR5) HDGNR10 polypeptide, useful for preventing or treating autoimmune  
 PT diseases e.g. rheumatoid arthritis, hyperproliferative disorders and  
 PT neurodegenerative disorders.  
 XX  
 PS Example 40; Page 504-505; 518pp; English.  
 XX  
 CC The invention relates to human G-protein chemokine receptor (CCR5)  
 CC HDGNR10 polypeptides and polynucleotides. CCR5 HDGNR10 antibodies are  
 CC useful for treating, preventing or ameliorating a disease or disorder  
 CC associated with inflammation, defective or aberrant chemotaxis of immune  
 CC cells, HIV infection (such as Pneumocystis carinii pneumonia or Kaposi's  
 CC sarcoma) or defective or aberrant T-cell antigen presenting cell  
 CC interaction. The disease or disorder may also be an infectious disease  
 CC (e.g. a viral infection such as an early stage HIV infection, a  
 CC cytomegalovirus infection, or a poxvirus infection), an autoimmune  
 CC disease (e.g. rheumatoid arthritis) or a neurodegenerative disorder. The  
 CC disease or disorder may be associated with aberrant CCR5 expression, lack  
 CC of CCR5 function, aberrant CCR5 ligand expression, or lack of CCR5 ligand  
 CC function. CCR5 HDGNR10 protein is used as a food additive or preservative  
 CC to increase or decrease storage capabilities. CCR5 HDGNR10 DNA are useful  
 CC for chromosome identification and in gene therapy. CCR5 HDGNR10 DNA,  
 CC protein, antibodies, agonists and antagonists are also useful in the  
 CC diagnosis, treatment and prevention of cancer (breast, ovary, adrenal  
 CC gland, bone, bone marrow, gastrointestinal tract, liver, lung,  
 CC urogenital); immune disorders (Addison's disease, allergies, autoimmune  
 CC haemolytic anaemia, autoimmune thyroiditis, diabetes mellitus, Crohn's  
 CC disease, multiple sclerosis, rheumatoid arthritis and ulcerative colitis)  
 CC ; cardiovascular disorders (myocardial ischaemias) and wound healing. The  
 CC present sequence is human CCR5 HDGNR10 protein  
 XX  
 SQ Sequence 352 AA;

Query Match 47.6%; Score 379; DB 4; Length 352;  
 Best Local Similarity 35.0%; Pred. No. 3.2e-31;  
 Matches 99; Conservative 7; Mismatches 27; Indels 150; Gaps 7;

QY 1 MDYQVSSPIYDINYYTSEPCQKINVKQIAA----- 30  
 ||||||||||||||||||  
 Db 1 MDYQVSSPIYDINYYTSEPCQKINVKQIAARLLPPLYSLVFIFGFVGNMLVILILINCKR 60  
 QY 31 -----YKCGLC-----AAAQWDFGNTMCQHQRVHGHHSYKCG--- 64

Db	61	LKSMTDIYLLNLAISDLFFLLTVPFWAHYAAAQWDFGNTMC--QLLTGLYFIGFFSGIFF	118
Qy	65	-----LCTRSQKEGLHYTC	78
Db	119	IILLTIDRYLAVVHAVFALKARTVTFGVVTSVITWVAVFASLPGLIFTRSQKEGLHYTC	178
Qy	79	SSHFPYSQYQFWKNFQTLKI-----HQRVHGG-----	105
Db	179	SSHFPYSQYQFWKNFQTLKIVILGLVLP LLVMVICYSGILKTLLRCRNEKKRHRVRLIF	238
Qy	106	-----GGSYKCGLC-----QEFG LNNC SSSNRLDGHQRV	135
Db	239	TIMIVYFLFWAPYNIVLLNTFQEFG LNNC SSSNRLDQAMQV	281

RESULT 8

AAG80111

ID AAG80111 standard; protein; 352 AA.  
 XX  
 AC AAG80111;  
 XX  
 DT 17-JAN-2002 (first entry)  
 XX  
 DE Human CCR5 protein.  
 XX  
 KW Chemokine; tumour diagnosis; colorectal; prostatic; organ rejection;  
 KW inflammation; autoimmune disease; metastasis; bronchial asthma; lupus;  
 KW chronic bowel inflammation; rheumatoid arthritis; cytostatic;  
 KW antiinflammatory; antiasthmatic; immunosuppressive; dermatological;  
 KW antirheumatic; antiarthritic.  
 XX  
 OS Homo sapiens.  
 XX  
 PN WO200172830-A2.  
 XX  
 PD 04-OCT-2001.  
 XX  
 PF 02-APR-2001; 2001WO-EP003708.  
 XX  
 PR 31-MAR-2000; 2000DE-01016013.  
 XX  
 PA (IPFP-) IPF PHARM GMBH.  
 PA (FORS/) FORSSMANN U.  
 XX  
 PI Forssmann W, Adermann K, Heitland A, Spodsberg N;  
 XX  
 DR WPI; 2001-626256/72.  
 XX  
 PT Diagnostic agent containing two or more receptor-specific ligands, useful  
 PT for detecting tumors, inflammation etc., also therapeutic use of ligand  
 PT inhibitors.  
 XX  
 PS Disclosure; Page 10; 26pp; German.  
 XX  
 CC This invention describes a novel diagnostic agent (A) comprising at least  
 CC two different ligands (I) for receptors (II) that are implicated in

CC disease. (A) are used for the diagnosis of tumors (especially colorectal  
 CC or prostatic), organ rejection, inflammation and autoimmune diseases.  
 CC Also inhibitors of (I) are used therapeutically against tumors (and their  
 CC metastases), inflammation (particularly bronchial asthma or chronic bowel  
 CC inflammation), or autoimmune diseases (rheumatoid arthritis or lupus),  
 CC where the (cardio)vascular, lymphatic, respiratory, nervous, digestive,  
 CC endocrine, motor or urogenital systems or skin are affected, and bone  
 CC marrow diseases. The products of the invention are chemokine derivatives  
 CC which have cytostatic, antiinflammatory, antiasthmatic,  
 CC immunosuppressive, dermatological, antirheumatic, antiarthritic.  
 CC Chemokines act on specific tumor and inflammatory cells through a  
 CC constellation of chemokine receptors (CR), which control migration and  
 CC proliferation of these cells. AAG80045-AAG80128 represent human chemokine  
 CC fragments used to illustrate the method of the invention  
 XX

SQ Sequence 352 AA;

Query Match 47.6%; Score 379; DB 4; Length 352;  
 Best Local Similarity 35.0%; Pred. No. 3.2e-31;  
 Matches 99; Conservative 7; Mismatches 27; Indels 150; Gaps 7;

QY	1 MDYQVSSPIYDINYYTSEPCQKINVKQIAA-----	30
Db	1 MDYQVSSPIYDINYYTSEPCQKINVKQIAARLLPPLYSLVFIFGFVGNMLVILILINCKR	60
QY	31 -----YKCGLC-----AAAQWDFGNTMCQHQRVHGHHHSYKCG---	64
	:   :	
Db	61 LKSMTDIYLLNLAISDLFFLLTVPFWAHYAAAQWDFGNTMC--QLLTGLYFIGFFSGIFF	118
QY	65 -----LCTRSQKEGLHYTC	78
	:	
Db	119 IILLTIDRYLAVVHAVFALKARTVTFGVVTSVITWVVAVFASLPGIIFTRSQKEGLHYTC	178
QY	79 SSHFPYSQYQFWKNFQTLKI-----HQRVHGG-----	105
	:	
Db	179 SSHFPYSQYQFWKNFQTLKIVILGLVLP LLVMVICYS GILKTLLRCRNEKKRHRVRLIF	238
QY	106 -----GGSYKCGLC----QEFFGLNNC SSSNRLDGHQRV	135
	:	
Db	239 TIMIVYFLFWAPYNIVLLLNTFQEFFGLNNC SSSNRLDQAMQV	281

RESULT 9

AAE04321

ID AAE04321 standard; protein; 352 AA.

XX

AC AAE04321;

XX

DT 04-SEP-2001 (first entry)

XX

DE Human chemokine receptor (CKR), CC-CKR-5 related protein #2.

XX

KW Human; transformed mammalian cell; CD4; reporter gene; translocation;

KW

KW human immuno deficiency virus; HIV; long terminal repeat; LTR; therapy;

KW

KW chemokine receptor; CKR; cellular dysfunction; HIV infection; cofactor;

XX

XX CC-CKR-5; envelope glycoprotein; anti-HIV.

OS Homo sapiens.  
 XX  
 PN US6258527-B1.  
 XX  
 PD 10-JUL-2001.  
 XX  
 PF 21-MAY-1997; 97US-00861105.  
 XX  
 PR 20-MAY-1996; 96US-0017157P.  
 PR 19-JUN-1996; 96US-0020043P.  
 PR 19-MAY-1997; 97US-00858660.  
 XX  
 PA (AARO-) AARON DIAMOND AIDS RES CENT.  
 PA (UYN Y ) UNIV NEW YORK STATE.  
 XX  
 PI Littman DR, Deng H, Ellmeier W, Landau NR, Liu R;  
 XX  
 DR WPI; 2001-417127/44.  
 DR N-PSDB; AAD08577.  
 XX  
 PT Transformed mammalian cell (I) that contains a CD4 gene, reporter gene  
 PT and HIV LTR for identification of drugs and antibodies for treatment of  
 PT HIV.  
 XX  
 PS Disclosure; Col 47-50; 37pp; English.  
 XX  
 CC The present invention relates to a transformed mammalian cell that  
 CC contains a gene encoding CD4, a construct encoding a reporter gene under  
 CC the regulation of an human immuno deficiency virus (HIV) long terminal  
 CC repeat (LTR) and that has been transduced with a vector encoding a human  
 CC chemokine receptor (CKR) where the CD4 and the CKR are present on the  
 CC cell surface of transformed mammalian cell. The invention is useful for  
 CC identifying drugs or antibodies that interfere with the translocation of  
 CC HIV into transformed mammalian cell or for identifying a human chemokine  
 CC receptor that facilitates the infection of a particular HIV strain into  
 CC the transformed mammalian cell. Compounds identified can be used to treat  
 CC cellular dysfunction and to prevent or combat HIV infection. The present  
 CC sequence is a human chemokine receptor (CKR), CC-CKR-5 related protein.  
 CC CC-CKR-5 is the principal cofactor for entry mediated by the envelope  
 CC glycoproteins of primary macrophage-tropic strains of HIV-1  
 XX  
 SQ Sequence 352 AA;

Query Match 47.6%; Score 379; DB 4; Length 352;  
 Best Local Similarity 35.0%; Pred. No. 3.2e-31;  
 Matches 99; Conservative 7; Mismatches 27; Indels 150; Gaps 7;

```

QY      1 MDYQVSSPIYDINYYTSEPCQKINVKQIAA----- 30
          |||
Db      1 MDYQVSSPIYDINYYTSEPCQKINVKQIAARLLPPLYSLVFIFGFVGNMLVILILINCKR 60
QY      31 -----YKCGLC-----AAAQWDFGNTMCQHQRVHGHHSYKCG--- 64
          |
Db      61 LKSMTDIYLLNLAISDLFFLLTVPFWAHYAAAQWDFGNTMC--QLLTGLYFIGFFSGIFF 118
QY      65 -----LCTRSQKEGLHYTC 78
          : |||

```

Db 119 IILLTIDRYLAVVHAVFALKARTVTFGVVTSVITWVAVFASLPGLIIFTRSQKEGLHYTC 178

Qy 79 SSHFPYSQYQFWKNFQTLKI-----HQRVHGG----- 105  
 |||||

Db 179 SSHFPYSQYQFWKNFQTLKIVILGLVLPLLMVICYSGILKTLLRCRNEKKRHRAVRLIF 238  
 :: |

Qy 106 -----GGSYKCGLC----QEFFGLNNCSSSNRLDGHQRV 135  
 | | |||||

Db 239 TIMIVYFLFWAPYNIVLLNTFQEFFGLNNCSSSNRLDQAMQV 281  
 :|

RESULT 10

AAE07039

ID AAE07039 standard; protein; 352 AA.

XX

AC AAE07039;

XX

DT 16-OCT-2001 (first entry)

XX

DE Human G-protein chemokine receptor (CCR5) HDGNR10 protein #2.

XX

KW Human; G-protein chemokine receptor; CCR5; HDGNR10; inflammation; HIV;  
 KW human immunodeficiency virus; antimicrobial; vasodilator; vulnerary;  
 KW cytostatic; immunosuppressive; nootropic; neuroprotective; gene therapy;  
 KW neurodegenerative disorder; Kaposi's sarcoma; autoimmune disease;  
 KW rheumatoid arthritis; cancer; breast; ovary; adrenal gland; bone marrow;  
 KW gastrointestinal tract; lung; liver; immune disorder; Addison's disease;  
 KW haemolytic anaemia; autoimmune thyroiditis; diabetes mellitus; allergy;  
 KW multiple sclerosis; ulcerative colitis; Crohn's disease; wound healing;  
 KW cardiovascular disorder; myocardial ischaemia.

XX

OS Homo sapiens.

XX

PN WO200158915-A2.

XX

PD 16-AUG-2001.

XX

PF 09-FEB-2001; 2001WO-US004152.

XX

PR 09-FEB-2000; 2000US-0181258P.

PR 09-MAR-2000; 2000US-0187999P.

PR 22-SEP-2000; 2000US-0234336P.

XX

PA (HUMA-) HUMAN GENOME SCI INC.

XX

PI Rosen CA, Roschke V, Li Y, Ruben SM;

XX

DR WPI; 2001-488965/53.

DR N-PSDB; AAD13198.

XX

PT Isolated nucleic acid encoding a human G-protein chemokine receptor  
 PT (CCR5) HDGNR10 polypeptide, useful for preventing or treating autoimmune  
 PT diseases e.g. rheumatoid arthritis, hyperproliferative disorders and  
 PT neurodegenerative disorders.

XX

PS Example 40; Page 486-487; 495pp; English.

XX

CC The present sequence is human G-protein chemokine receptor (CCR5) HDGNR10  
 CC protein. CCR5 HDGNR10 antibodies are useful for treating, preventing or  
 CC ameliorating a disease or disorder associated with inflammation,  
 CC defective or aberrant chemotaxis of immune cells, HIV infection (such as  
 CC *Pneumocystis carinii* pneumonia or Kaposi's sarcoma) or defective or  
 CC aberrant T-cell antigen presenting cell interaction. The disease or  
 CC disorder may also be an infectious disease (e.g. a viral infection such  
 CC as an early stage HIV infection, a cytomegalovirus infection, or a  
 CC poxvirus infection), an autoimmune disease (e.g. rheumatoid arthritis) or  
 CC a neurodegenerative disorder. The disease or disorder may be associated  
 CC with aberrant CCR5 expression, lack of CCR5 function, aberrant CCR5  
 CC ligand expression, or lack of CCR5 ligand function. CCR5 HDGNR10 protein  
 CC is used as a food additive or preservative to increase or decrease  
 CC storage capabilities. CCR5 HDGNR10 DNA are useful for chromosome  
 CC identification and in gene therapy. CCR5 HDGNR10 DNA, protein,  
 CC antibodies, agonists and antagonists are also useful in the diagnosis,  
 CC treatment and prevention of cancer (breast, ovary, adrenal gland, bone,  
 CC bone marrow, gastrointestinal tract, liver, lung, urogenital); immune  
 CC disorders (Addison's disease, allergies, autoimmune haemolytic anaemia,  
 CC autoimmune thyroiditis, diabetes mellitus, Crohn's disease, multiple  
 CC sclerosis, rheumatoid arthritis and ulcerative colitis); cardiovascular  
 CC disorders (myocardial ischaemias) and wound healing  
 XX  
 SQ Sequence 352 AA;

Query Match 47.6%; Score 379; DB 4; Length 352;  
 Best Local Similarity 35.0%; Pred. No. 3.2e-31;  
 Matches 99; Conservative 7; Mismatches 27; Indels 150; Gaps 7;

```

QY      1 MDYQVSSPIYDINYYTSEPCQKINVKQIAA----- 30
          |||
Db      1 MDYQVSSPIYDINYYTSEPCQKINVKQIAARLLPPLYSLVFIFGFVGNMLVILILINCKR 60
QY      31 -----YKCGLC-----AAAQWDFGNTMCQHQRVHGHHSYKCG--- 64
          | |
Db      61 LKSMTDIYLLNLAIISDLFFLLTVPFWAHYAAAQWDFGNTMC--QLLTGLYFIGFFSGIFF 118
QY      65 -----LCTRSQKEGLHYTC 78
          : |||
Db      119 IILLTIDRYLAVVHAVFALKARTVTFGVVTSVITWVVAVFASLPGIIFTRSQKEGLHYTC 178
QY      79 SSHFPYSQYQFWKNFQTLKI-----HQRVHGG----- 105
          |||
Db      179 SSHFPYSQYQFWKNFQTLKIVILGLVLPPLVMVICYSGILKTLLRCRNEKKRHRAVRLIF 238
QY      106 -----GGSYKCGLC----QEFFGLNNCSSSNRLDGHQRV 135
          | |
Db      239 TIMIVYFLFWAPYNIVLLLNTFQEFFGLNNCSSSNRLDQAMQV 281

```

# RESULT 11

AAB46858

ID AAB46858 standard; protein; 352 AA.

XX

AC AAB46858;

XX

DT 16-AUG-2001 (revised)

DT 02-AUG-2001 (revised)  
 DT 04-MAY-2001 (first entry)  
 XX  
 DE Human HDGNR10 protein.  
 XX  
 KW HDGNR10; human; G-protein chemokine receptor; antiinflammatory;  
 KW immunomodulatory; anticoagulant; antiallergic; immunosuppressive;  
 KW cytostatic; antiparasitic; antipsoriatic; antirheumatic; antiarthritic;  
 KW vasotropic; gene therapy; haematopoiesis; wound healing; coagulation;  
 KW angiogenesis; solid tumour; infection; leukemia; growth factor activity;  
 KW T-cell mediated autoimmune disease; psoriasis; allergy; atherogenesis;  
 KW anaphylaxis; malignancy; inflammation; histamine; IgE; silicosis; shock;  
 KW immunoglobulin E-mediated allergic reaction; rheumatoid arthritis;  
 KW prostaglandin-independent fever; bone marrow failure; sarcoidosis;  
 KW hyper-eosinophilic syndrome; vulnerary.  
 XX  
 OS Homo sapiens.  
 XX  
 PN US2001000241-A1.  
 XX  
 PD 12-APR-2001.  
 XX  
 PF 29-NOV-2000; 2000US-00725285.  
 XX  
 PR 06-JUN-1995; 95US-00466343.  
 PR 18-NOV-1998; 98US-00195662.  
 PR 25-JUN-1999; 99US-00339912.  
 XX  
 PA (LIYY/) LI Y.  
 PA (RUBE/) RUBEN S M.  
 XX  
 PI Li Y, Ruben SM;  
 XX  
 DR WPI; 2001-226317/23.  
 DR N-PSDB; AAF26390.  
 XX  
 PT New human G-protein chemokine receptor polypeptides and polynucleotides,  
 PT useful for identifying (ant)agonists to the G-protein chemokine receptor.  
 XX  
 PS Claim 1a; Page 15; 22pp; English.  
 XX  
 CC This invention describes a novel receptor polypeptide (I) selected from  
 CC (i) a fully defined 329 amino acid sequence (II) fully disclosed in the  
 CC specification; and (ii) a polypeptide encoded by the cDNA contained in a  
 CC plasmid, and fragments, analogs and derivatives of the polypeptide. The  
 CC products of the invention have antiinflammatory, immunomodulatory,  
 CC anticoagulant, antiallergic, immunosuppressive, vulnerary, cytostatic,  
 CC antiparasitic, antipsoriatic, antirheumatic, antiarthritic and vasotropic  
 CC activity and can be used for gene therapy. The G-protein chemokine  
 CC receptors, HDGNR10, (I) are useful for screening for compounds which  
 CC activate or inhibit activation of (I). The products of the invention can  
 CC also be used for stimulating haematopoiesis, wound healing, coagulation,  
 CC angiogenesis, treating solid tumours, chronic infections, leukemia, T-  
 CC cell mediated autoimmune diseases, parasitic infections, psoriasis, and  
 CC stimulating growth factor activity. HDGNR10 is useful for treating  
 CC allergy, atherogenesis, anaphylaxis, malignancy, chronic and acute  
 CC inflammation, histamine and immunoglobulin E (IgE)-mediated allergic





XX  
 PI Lehmann-Bruinsma K, Liaw CW, Lin I;  
 XX  
 DR WPI; 2001-648759/74.  
 DR N-PSDB; ABI97978.  
 XX  
 PT Identifying agonists of G protein-coupled receptors (GPCRs) for use in  
 PT disease treatment, comprises contacting candidate compounds with versions  
 PT of GPCRs.  
 XX  
 PS Claim 1; Page 277-278; 394pp; English.  
 XX  
 CC The invention relates to G protein-coupled receptors (GPCRs) for which  
 CC the endogenous ligand has been identified. Non-endogenous constitutively  
 CC activated versions of known GPCRs are used in the invention for the  
 CC direct identification of candidate compounds as receptor agonists,  
 CC inverse agonists or partial agonists. Such agonists are useful as  
 CC therapeutic agents for diseases or disorders associated with GPCRs. The  
 CC present sequence is a non-endogenous version of a known human GPCR  
 XX  
 SQ Sequence 352 AA;

Query Match 47.6%; Score 379; DB 4; Length 352;  
 Best Local Similarity 35.0%; Pred. No. 3.2e-31;  
 Matches 99; Conservative 7; Mismatches 27; Indels 150; Gaps 7;

QY 1 MDYQVSSPIYDINYYTSEPCQKINVKQIAA----- 30  
 ||||||||||||||||||||||||||||  
 Db 1 MDYQVSSPIYDINYYTSEPCQKINVKQIAARLLPPLYSLVFIFGFVGNMLVILILINCKR 60  
 QY 31 -----YKCGLC-----AAAQWDFGNTMCQHQRVHGHHHSYKCG--- 64  
 | | ||||||||| | : | : : |  
 Db 61 LKSMTDIYLLNLAISDLFFLLTVPFWAHYAAAQWDFGNTMC--QLLTGLYFIGFFSGIFF 118  
 QY 65 -----LCTRSQKEGLHYTC 78  
 : |||||||||  
 Db 119 IILLTIDRYLAVVHAVFALKARTVTFGVVTSVITWVVAVFASLPGIIFTRSQKEGLHYTC 178  
 QY 79 SSHFPYSQYQFWKNFQTLKI-----HQRVHGG----- 105  
 ||||||||||||||||| : : |  
 Db 179 SSHFPYSQYQFWKNFQTLKIVILGLVLPLLMVICYSGILKTLLRCRNEKKRHRARLIF 238  
 QY 106 -----GGSYKCGLC----QEFGFLNNCSSSNRLDGHQRV 135  
 | | ||||||||||||| : |  
 Db 239 TIMIVYFLFWAPYNIVLLLNTFQEFGFLNNCSSSNRLDQAMQV 281

RESULT 13

AAB83354

ID AAB83354 standard; protein; 352 AA.

XX

AC AAB83354;

XX

DT 09-OCT-2001 (first entry)

XX

DE Human CCR5 protein sequence.

XX

KW Chemotactic chemokine receptor 5; gp120; CD4; therapy; HIV; CCR5;  
 KW human immunodeficiency virus; anti-inflammatory disease; human.  
 XX  
 OS Homo sapiens.  
 XX  
 PN EP1118858-A2.  
 XX  
 PD 25-JUL-2001.  
 XX  
 PF 03-JAN-2001; 2001EP-00300020.  
 XX  
 PR 12-JAN-2000; 2000GB-00000659.  
 PR 12-JAN-2000; 2000GB-00000661.  
 PR 12-JAN-2000; 2000GB-00000663.  
 XX  
 PA (PFIZ ) PFIZER LTD.  
 PA (PFIZ ) PFIZER INC.  
 XX  
 PI Dobbs S, Perros M, Rickett GA;  
 XX  
 DR WPI; 2001-477088/52.  
 DR N-PSDB; AAF87099.  
 XX  
 PT Determining if an agent can modulate CCR5-gp120 interaction, comprises  
 PT incubating the agent with CCR5 and gp120 and determining if the agent  
 PT modulates the interaction.  
 XX  
 PS Claim 1; Page 110; 113pp; English.  
 XX  
 CC This sequence represents the human CCR5 protein sequence. The invention  
 CC relates to a method for determining whether an agent is capable of  
 CC modulating the interaction of chemotactic chemokine receptor 5 (CCR5)  
 CC with gp120, comprising incubating the agent with CCR5 and gp120 and  
 CC determining whether the agent modulates the interaction, where gp120 is  
 CC associated with CD4, and where the interaction is a low affinity binding.  
 CC The method is used to identify an agent capable of modulating the  
 CC interaction of CCR5 with gp120. An agent identified by the method is used  
 CC to prepare a pharmaceutical composition for the treatment of a disease or  
 CC condition associated with CCR5 and gp120 interaction, to treat a subject  
 CC with a disease or condition associated with CCR5 and gp120 interaction,  
 CC and for preparing a pharmaceutical for treating human immunodeficiency  
 CC virus (HIV). It can also be treat anti-inflammatory diseases. The method  
 CC is commercially useful, amenable to high throughput screening, and  
 CC detects interaction of gp120 with cells expressing only CCR5  
 XX  
 SQ Sequence 352 AA;

Query Match 47.6%; Score 379; DB 4; Length 352;  
 Best Local Similarity 35.0%; Pred. No. 3.2e-31;  
 Matches 99; Conservative 7; Mismatches 27; Indels 150; Gaps 7;

QY 1 MDYQVSSPIYDINYYTSEPCQKINVKQIAA----- 30  
 ||||||||||||||||||||  
 Db 1 MDYQVSSPIYDINYYTSEPCQKINVKQIAARLLPPLYSLVFIFGFVGNMLVILILINCKR 60  
 QY 31 -----YKCGLC-----AAAQWDFGNTMCQHQRVHGHHSYKCG--- 64  
 | | |||||||| | : | : : |

Db 61 LKSMTDIYLLNLAISDLFFLLTVPFWAHYAAQWDFGNTMC--QLLTGLYFIGFFSGIFF 118  
 QY 65 -----LCTRSQKEGLHYTC 78  
 Db 119 IILLTIDRYLAVVHAVFALKARTVTFGVVTSVITWVAVFASLPGIIFTRSQKEGLHYTC 178  
 QY 79 SSHFPYSQYQFWKNFQTLKI-----HQRVHGG----- 105  
 Db 179 SSHFPYSQYQFWKNFQTLKIVILGLVLP LLVMVICYSGILKTLLRCRNEKKRHRVRLIF 238  
 QY 106 -----GGSYKCGLC----QEFFGLNNC SSSNRLDGHQRV 135  
 Db 239 TIMIVYFLFWAPYNIVLLNTFQEFFGLNNC SSSNRLDQAMQV 281

RESULT 14

AAB82948

ID AAB82948 standard; protein; 352 AA.  
 XX  
 AC AAB82948;  
 XX  
 DT 21-DEC-2001 (first entry)  
 XX  
 DE Human HIV-1 co-receptor CCR5.  
 XX  
 KW CCR5; chemokine; co-receptor; human immunodeficiency virus type 1; HIV-1;  
 KW infection; therapy; vaccine; anti-HIV-1.  
 XX  
 OS Homo sapiens.  
 XX  
 FH Key Location/Qualifiers  
 FT Binding-site 2. .18  
 FT /note= "binds to HIV-1 gp120"  
 XX  
 PN WO200164710-A2.  
 XX  
 PD 07-SEP-2001.  
 XX  
 PF 28-FEB-2001; 2001WO-US006699.  
 XX  
 PR 29-FEB-2000; 2000US-0185667P.  
 PR 19-MAY-2000; 2000US-0205839P.  
 PR 07-FEB-2001; 2001US-0267231P.  
 XX  
 PA (PROG-) PROGENICS PHARM INC.  
 PA (AARO-) AARON DIAMOND AIDS RES CENT.  
 XX  
 PI Dragic T, Olson WC;  
 XX  
 DR WPI; 2001-611273/70.  
 DR N-PSDB; AAH26903.  
 XX  
 PT Novel compounds comprising specific amino acids within CCR5 (HIV 1 co-  
 PT receptor) amino terminal domain including negatively charged and two  
 PT sulfated tyrosine residues is useful for treating HIV infection in  
 PT humans.  
 XX

PS Claim 1; Page 30; 163pp; English.

XX

CC The present sequence is that of human HIV-1 co-receptor CCR5. Amino acids  
CC 2-18 in the N-terminal region of CCR5 comprise an HIV-1 gp120-binding  
CC site that determines the specificity of the interaction between CCR5 and  
CC HIV-1 gp120. Post-translational sulfation of the tyrosine residues in the  
CC CCR5 N-terminus is required for gp120 binding and may critically modulate  
CC the susceptibility of target cells to HIV-1 infection in vivo. The  
CC invention provides claimed sulfated peptides (see AAB82947) that are  
CC based on the CCR5 N-terminal region and which are effective for  
CC inhibiting HIV-1 binding to CCR5. These peptides are used in claimed  
CC methods of inhibiting HIV infection of CD4+ cells, of preventing CD4+  
CC cells from becoming infected with HIV, of treating a subject whose CD4+  
CC cells are infected with HIV, and of identifying an agent which inhibits  
CC binding of a CCR5 ligand to a CCR5 receptor. The methods may be carried  
CC out in a subject, especially a human, infected (therapeutic method), not  
CC infected with HIV (prophylactic method), or in a subject who is not  
CC infected with, but has been exposed to, HIV

XX

SQ Sequence 352 AA;

Query Match 47.6%; Score 379; DB 4; Length 352;

Best Local Similarity 35.0%; Pred. No. 3.2e-31;

Matches 99; Conservative 7; Mismatches 27; Indels 150; Gaps 7;

```
QY      1 MDYQVSSPIYDINYYTSEPCQKINVKQIAA----- 30
          |||
Db      1 MDYQVSSPIYDINYYTSEPCQKINVKQIAARLLPPLYSLVFIFGFVGNMLVILILINCKR 60
          |||

QY      31 -----YKCGLC-----AAAQWDFGNTMCQHQRVHGHHSYKCG--- 64
          | | |
Db      61 LKSMTDIYLLNLAISDLFFLLTVPFWAHYAAAQWDFGNTMC--QLLTGLYFIGFFSGIFF 118
          | | |

QY      65 -----LCTRSQKEGLHYTC 78
          : |||
Db      119 IILLTIDRYLAVVHAVFALKARTVTFGVVTSVITWVVAVFASLPGIIFTRSQKEGLHYTC 178
          : |||

QY      79 SSHFPYSQYQFWKNFQTLKI-----HQRVHGG----- 105
          |||
Db      179 SSHFPYSQYQFWKNFQTLKIVILGLVLP LLVMVICYSGILKTLLRCRNEKKRHRVRLIF 238
          : |

QY      106 -----GGSYKCGLC----QEFGFLNNCSSSNRLDGHQRV 135
          | | |
Db      239 TIMIVYFLFWAPYNIVLLLNTFQEFFGLNNCSSSNRLDQAMQV 281
          | | |
```

RESULT 15

AAU97152

ID AAU97152 standard; protein; 352 AA.

XX

AC AAU97152;

XX

DT 13-AUG-2002 (first entry)

XX

DE Human G-protein chemokine receptor (CCR5) HDG NR10 #2.

XX

KW Human; G-protein chemokine receptor; CCR5; HDG NR10; inflammation;

KW immune cell chemotaxis; autoimmune disease; rheumatoid arthritis;  
 KW neurodegeneration; viral infection; Kaposi sarcoma; cancer;  
 KW hyperproliferative disease; neurological disease; receptor.  
 XX  
 OS Homo sapiens.  
 XX  
 PN US2002048786-A1.  
 XX  
 PD 25-APR-2002.  
 XX  
 PF 09-FEB-2001; 2001US-00779879.  
 XX  
 PR 09-FEB-2000; 2000US-0181258P.  
 PR 09-MAR-2000; 2000US-0187999P.  
 PR 22-SEP-2000; 2000US-0234336P.  
 XX  
 PA (ROSE/) ROSEN C A.  
 PA (ROSC/) ROSCHKE V.  
 PA (LIYY/) LI Y.  
 PA (RUBE/) RUBEN S M.  
 XX  
 PI Rosen CA, Roschke V, Li Y, Ruben SM;  
 XX  
 DR WPI; 2002-434754/46.  
 DR N-PSDB; ABK51870.  
 XX  
 PT New nucleic acid encoding an antibody specific for the G-protein  
 PT chemokine receptor CCR5, useful for treatment and diagnosis of e.g.  
 PT inflammation.  
 XX  
 PS Disclosure; Page 165-166; 180pp; English.  
 XX  
 CC The present invention relates to the isolation of a novel human G-protein  
 CC chemokine receptor (CCR5) designated HDGNR10, and polynucleotide  
 CC sequences encoding it. The invention also describes antibodies that bind  
 CC human G-protein chemokine receptor (CCR5) HDGNR10 and polynucleotide  
 CC sequences encoding the antibodies. The antibodies are useful for treating  
 CC or preventing inflammation, defective or aberrant chemotaxis of immune  
 CC cells and T-cell/antigen-presenting cell interactions, infections and  
 CC autoimmune diseases, rheumatoid arthritis, neurodegeneration, viral  
 CC infections (especially early-stage human immune deficiency virus (HIV),  
 CC cytomegalovirus or pox virus infections, Kaposi sarcoma, or conditions  
 CC associated with aberrant or deficient expression of the CCR5 receptor or  
 CC its ligands. The antibodies are also useful to determine CCR5 expression,  
 CC e.g. for diagnosis, prognosis and monitoring of cancer and other  
 CC hyperproliferative diseases. The polynucleotide sequences encoding human  
 CC G-protein chemokine receptor (CCR5) HDGNR10 can be used to produce the  
 CC recombinant receptor, and in the treatment of a wide range of diseases  
 CC such as infectious diseases (e.g. influenza), neurological diseases (e.g.  
 CC Parkinson's disease), and hyperproliferative disorders (e.g. cancer). The  
 CC present sequence represents human G-protein chemokine receptor (CCR5)  
 CC HDGNR10 #2  
 XX  
 SQ Sequence 352 AA;

Query Match 47.6%; Score 379; DB 5; Length 352;  
 Best Local Similarity 35.0%; Pred. No. 3.2e-31;

Matches 99; Conservative 7; Mismatches 27; Indels 150; Gaps 7;

```
Qy      1 MDYQVSSPIYDINYTSEPCQKINVKQIAA----- 30
          ||||||||||||||||||||||||||||
Db      1 MDYQVSSPIYDINYTSEPCQKINVKQIAARLLPPLYSLVFIFGFVGNMLVILILINCKR 60

Qy     31 -----YKCGLC-----AAAQWDFGNTMCQHQRVHGHHSYKCG--- 64
          |  |                      |||||  | : | : |
Db     61 LKSMTDIYLLNLAISDLFFLLTVPFWAHYAAAQWDFGNTMC--QLLTGLYFIGFFSGIFF 118

Qy     65 -----LCTRSQKEGLHYTC 78
          : |||||||
Db    119 IILLTIDRYLAVVHAVFALKARTVTFGVVTSVITWVVAVFASLPGIIFTRSQKEGLHYTC 178

Qy     79 SSHFPYSQYQFWKNFQTLKI-----HQRVHGG----- 105
          ||||||||||||||||  :: |
Db    179 SSHFPYSQYQFWKNFQTLKIVILGLVLP LLVMVICYSGILKTLLRCRNEKKRHRVRLIF 238

Qy    106 -----GGSYKCGLC----QEFFGLNNCSSSNRLDGHQRV 135
          |  |  ||||||||||||  :|
Db    239 TIMIVYFLFWAPYNIVLLLNTFQEFFGLNNCSSSNRLDQAMQV 281
```

Search completed: September 17, 2004, 16:45:08  
Job time : 83.461 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: September 17, 2004, 16:42:06 ; Search time 21.5186 Seconds  
(without alignments)  
331.080 Million cell updates/sec

Title: US-10-057-890A-10  
Perfect score: 797  
Sequence: 1 MDYQVSSPIYDINYTSEPC.....GLNNCSSSNRLDGHQRVHAA 138

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 389414 seqs, 51625971 residues

Total number of hits satisfying chosen parameters: 389414

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Issued\_Patents\_AA:\*  
1: /cgn2\_6/ptodata/2/iaa/5A\_COMB.pep:\*  
2: /cgn2\_6/ptodata/2/iaa/5B\_COMB.pep:\*  
3: /cgn2\_6/ptodata/2/iaa/6A\_COMB.pep:\*  
4: /cgn2\_6/ptodata/2/iaa/6B\_COMB.pep:\*  
5: /cgn2\_6/ptodata/2/iaa/PCTUS\_COMB.pep:\*  
6: /cgn2\_6/ptodata/2/iaa/backfiles1.pep:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

		%					
Result		Query					
No.	Score	Match	Length	DB	ID	Description	
1	379	47.6	352	3	US-09-087-232A-13	Sequence 13, Appl	
2	379	47.6	352	3	US-08-861-105-14	Sequence 14, Appl	
3	379	47.6	352	3	US-08-575-967A-2	Sequence 2, Appli	
4	379	47.6	352	4	US-08-833-752-5	Sequence 5, Appli	
5	379	47.6	352	4	US-09-502-783A-2	Sequence 2, Appli	
6	379	47.6	352	4	US-09-796-202-1	Sequence 1, Appli	
7	374	46.9	352	3	US-09-045-583-52	Sequence 52, Appl	
8	374	46.9	352	4	US-09-534-185-52	Sequence 52, Appl	
9	373	46.8	352	3	US-08-466-343D-2	Sequence 2, Appli	
10	359	45.0	352	4	US-09-517-605-5	Sequence 5, Appli	
11	272	34.1	354	4	US-08-724-984A-2	Sequence 2, Appli	

12	258	32.4	184	4	US-08-833-752-4	Sequence 4, Appli
13	258	32.4	215	3	US-09-087-232A-17	Sequence 17, Appl
14	258	32.4	215	4	US-08-833-752-6	Sequence 6, Appli
15	182.5	22.9	100	3	US-09-087-232A-15	Sequence 15, Appl
16	159	19.9	87	3	US-09-087-232A-18	Sequence 18, Appl
17	148.5	18.6	488	2	US-08-933-750C-17	Sequence 17, Appl
18	148.5	18.6	488	3	US-09-234-613-17	Sequence 17, Appl
19	128	16.1	54	4	US-08-833-752-11	Sequence 11, Appl
20	128	16.1	313	4	US-09-800-729-196	Sequence 196, App
21	128	16.1	326	4	US-09-800-729-195	Sequence 195, App
22	123.5	15.5	717	4	US-09-881-578A-2	Sequence 2, Appli
23	123	15.4	22	3	US-08-861-105-4	Sequence 4, Appli
24	121	15.2	518	4	US-09-881-578A-4	Sequence 4, Appli
25	121	15.2	711	2	US-08-820-170A-10	Sequence 10, Appl
26	121	15.2	711	3	US-09-055-699-10	Sequence 10, Appl
27	121	15.2	711	3	US-09-273-565-10	Sequence 10, Appl
28	121	15.2	711	4	US-09-565-538-10	Sequence 10, Appl
29	121	15.2	711	4	US-09-661-468-10	Sequence 10, Appl
30	121	15.2	711	4	US-09-976-165-10	Sequence 10, Appl
31	120.5	15.1	185	4	US-09-494-190-125	Sequence 125, App
32	119.5	15.0	185	4	US-09-494-190-126	Sequence 126, App
33	118.5	14.9	315	1	US-08-253-155A-34	Sequence 34, Appl
34	118.5	14.9	532	4	US-09-389-956-80	Sequence 80, Appl
35	117	14.7	457	4	US-09-389-956-68	Sequence 68, Appl
36	117	14.7	647	4	US-09-389-956-6	Sequence 6, Appli
37	114.5	14.4	543	4	US-09-362-123A-4	Sequence 4, Appli
38	114.5	14.4	675	1	US-08-317-522A-9	Sequence 9, Appli
39	114.5	14.4	675	1	US-08-439-818A-9	Sequence 9, Appli
40	114.5	14.4	675	2	US-08-751-965-9	Sequence 9, Appli
41	114.5	14.4	675	2	US-08-738-975-9	Sequence 9, Appli
42	114.5	14.4	675	2	US-08-728-626-9	Sequence 9, Appli
43	114.5	14.4	675	3	US-08-808-599A-9	Sequence 9, Appli
44	114	14.3	21	3	US-08-907-468-11	Sequence 11, Appl
45	113.5	14.2	462	3	US-08-486-099-117	Sequence 117, App

# ALIGNMENTS

## RESULT 1

US-09-087-232A-13

; Sequence 13, Application US/09087232A

; Patent No. 6153431

; GENERAL INFORMATION:

; APPLICANT: Quillent et al.

; TITLE OF INVENTION: HUMAN IMMUNODEFICIENCY VIRUS CO-RECEPTOR

; TITLE OF INVENTION: VARIANTS ASSOCIATED WITH RESISTANCE TO VIRUS INFECTION.

; NUMBER OF SEQUENCES: 23

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Baker & Botts, L.L.P. attn. Lisa Kole

; STREET: 30 Rockefeller Plaza

; CITY: New York

; STATE: New York

; COUNTRY: USA

; ZIP: 10112

; COMPUTER READABLE FORM:



```

; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30 (EPO)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/087,232A
; FILING DATE: 28 MAY 1998
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 60/048,057
; FILING DATE: 30 MAY 1997
; ATTORNEY/AGENT INFORMATION:
; NAME: KOLE, LISA B.
; REGISTRATION NUMBER: 35,225
; REFERENCE/DOCKET NUMBER: AP 31115
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 408-2628
; TELEFAX: (212) 765-2519
; INFORMATION FOR SEQ ID NO: 13:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 352 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-09-087-232A-13

```

```

Query Match          47.6%; Score 379; DB 3; Length 352;
Best Local Similarity 35.0%; Pred. No. 1.7e-34;
Matches 99; Conservative 7; Mismatches 27; Indels 150; Gaps 7;

```

```

QY      1 MDYQVSSPIYDINYYTSEPCQKINVKQIAA----- 30
          |||||||||||||||||||
Db      1 MDYQVSSPIYDINYYTSEPCQKINVKQIAARLLPPLYSLVFIFGFVGNMLVILILINCKR 60
QY      31 -----YKCGLC-----AAAQWDFGNTMCQHQRVHGHHSYKCG--- 64
          | | ||||||||| | : | : : |
Db      61 LKSMTDIYLLNLAISDLFFLLTVPFWAHYAAAQWDFGNTMC--QLLTGLYFIGFFSGIFF 118
QY      65 -----LCTRSQKEGLHYTC 78
          : |||||||||
Db      119 IILLTIDRYLAVVHAVFALKARTVTFGVVTSVITWVVAVFASLPGIIFTRSQKEGLHYTC 178
QY      79 SSHFPYSQYQFWKNFQTLKI-----HQRVHGG----- 105
          ||||||||||||||||| : : |
Db      179 SSHFPYSQYQFWKNFQTLKIVILGLVLPPLVMVICYSGILKTLLRCRNEKKRHRAVRLIF 238
QY      106 -----GGSYKCGLC----QEFFGLNNCSSSNRLDGHQRV 135
          | | ||||||||| :|
Db      239 TIMIVYFLWAPYNIVLLLNTFQEFFGLNNCSSSNRLDQAMQV 281

```

RESULT 2

US-08-861-105-14

```

; Sequence 14, Application US/08861105
; Patent No. 6258527
; GENERAL INFORMATION:
; APPLICANT: LITTMAN, DAN R.

```

```

; APPLICANT: DENG, HONGKUI
; APPLICANT: ELLMEIER, WILFRIED
; APPLICANT: LANDAU, NATHANIEL R.
; APPLICANT: LIU, RONG
; TITLE OF INVENTION: G-COUPLED RECEPTORS ASSOCIATED WITH
; TITLE OF INVENTION: MACROPHAGE-TROPHIC HIV, AND DIAGNOSTIC AND THERAPEUTIC
; TITLE OF INVENTION: USES THEREOF
; NUMBER OF SEQUENCES: 14
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: David A. Jackson, Esq.
; STREET: 411 Hackensack Ave, Continental Plaza, 4th
; STREET: Floor
; CITY: Hackensack
; STATE: New Jersey
; COUNTRY: USA
; ZIP: 07601
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/861,105
; FILING DATE:
; CLASSIFICATION: 436
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/666,020
; FILING DATE: 19-JUN-1996
; CLASSIFICATION: 436
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/227,319
; FILING DATE: 13-APR-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: Jackson Esq., David A.
; REGISTRATION NUMBER: 26,742
; REFERENCE/DOCKET NUMBER: 1049-1-004 N1
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 201-487-5800
; TELEFAX: 201-343-1684
; INFORMATION FOR SEQ ID NO: 14:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 352 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; HYPOTHETICAL: NO
; ORIGINAL SOURCE:
; ORGANISM: Homo sapiens
US-08-861-105-14

```

```

Query Match          47.6%; Score 379; DB 3; Length 352;
Best Local Similarity 35.0%; Pred. No. 1.7e-34;
Matches 99; Conservative 7; Mismatches 27; Indels 150; Gaps 7;

```

```

QY      1 MDYQVSSPIYDINYTSEPCQKINVKQIAA----- 30
      ||||||||||||||||||||||||||||

```

Db 1 MDYQVSSPIYDINYTTSEPCQKINVKQIAARLLPPLYSLVFIFGFVGNMLVILILINCKR 60  
 Qy 31 -----YKCGLC-----AAAQWDFGNTMCQHQRVHGHHSYKCG--- 64  
 | | | | | | | | | | : | : : |  
 Db 61 LKSMTDIYLLNLAISDLFFLLTVPFWAHYAAAQWDFGNTMC--QLLTGLYFIGFFSGIFF 118  
 Qy 65 -----LCTRSQKEGLHYTC 78  
 : | | | | | | | | | |  
 Db 119 IILLTIDRYLAVVHAVFALKARTVTFGVVTSVITWVAVFASLPGIIFTRSQKEGLHYTC 178  
 Qy 79 SSHFPYSQYQFWKNFQTLKI-----HQRVHGG----- 105  
 | | | | | | | | | | : : |  
 Db 179 SSHFPYSQYQFWKNFQTLKIVILGLVLP LLVMVICYSGILKTLLRCRNEKKRHRVRLIF 238  
 Qy 106 -----GGSYKCGLC-----QEFFGLNNCSSSNRLDGHQRV 135  
 | | | | | | | | | | : |  
 Db 239 TIMIVYFLFWAPYNIVLLLNTFQEFFGLNNCSSSNRLDQAMQV 281

RESULT 3

US-08-575-967A-2

; Sequence 2, Application US/08575967A

; Patent No. 6265184

; GENERAL INFORMATION:

; APPLICANT: Gray et al.

; TITLE OF INVENTION: Chemokine Receptor Materials and Methods

; NUMBER OF SEQUENCES: 16

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Marshall, O'Toole, Gerstein, Murray & Borun

; STREET: 6300 Sears Tower, 233 S. Wacker Drive

; CITY: Chicago

; STATE: Illinois

; COUNTRY: USA

; ZIP: 60606

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk

; COMPUTER: IBM PC compatible

; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: PatentIn Release #1.0, Version #1.30

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/08/575,967A

; FILING DATE:

; CLASSIFICATION: 435

; ATTORNEY/AGENT INFORMATION:

; NAME: No. 6265184and, Greta E.

; REGISTRATION NUMBER: 35,302

; REFERENCE/DOCKET NUMBER: 32918

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: 206-485-1900

; TELEFAX: 206-485-1662

; INFORMATION FOR SEQ ID NO: 2:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 352 amino acids

; TYPE: amino acid

; TOPOLOGY: linear

; MOLECULE TYPE: protein

; FEATURE:

; NAME/KEY: misc\_feature  
; OTHER INFORMATION: /= "88C amino acid sequence"  
US-08-575-967A-2

Query Match 47.6%; Score 379; DB 3; Length 352;  
Best Local Similarity 35.0%; Pred. No. 1.7e-34;  
Matches 99; Conservative 7; Mismatches 27; Indels 150; Gaps 7;

```
Qy      1 MDYQVSSPIYDINYYTSEPCQKINVKQIAA----- 30
          ||||||||||||||||||||||||||||
Db      1 MDYQVSSPIYDINYYTSEPCQKINVKQIAARLLPPLYSLVFIFGFVGNMLVILILINCKR 60

Qy     31 -----YKCGLC-----AAAQWDFGNTMCQHQRVHGHHHSYKCG--- 64
          |   |               ||||||||| | : | : : |
Db     61 LKSMTDIYLLNLAISDLFFLLTVPFWAHYAAAQWDFGNTMC--QLLTGLYFIGFFSGIFF 118

Qy     65 -----LCTRSQKEGLHYTC 78
          : |||||||||||
Db    119 IILLTIDRYLAVVHAVFALKARTVTFGVVT SVITWVAVFASLPGIIFTRSQKEGLHYTC 178

Qy     79 SSHFPYSQYQFWKNFQTLKI-----HQRVHGG----- 105
          |||||||||||||||||
Db    179 SSHFPYSQYQFWKNFQTLKIVILGLVLP LLVMVICYSGILKTL LRCRNEKKRHRAVRLIF 238

Qy    106 -----GGSYKCGLC----QEFG LNNC SSSNRLDGHQRV 135
          |   |   ||||||||||||| :|
Db    239 TIMIVYFLFWAPYNIVLLNLNTFQEFG LNNC SSSNRLDQAMQV 281
```

RESULT 4

US-08-833-752-5

; Sequence 5, Application US/08833752

; Patent No. 6448375

; GENERAL INFORMATION:

; APPLICANT: SAMSON, MICHEL

; APPLICANT: PARMENTIER, MARC

; APPLICANT: VASSART, GILBERT

; APPLICANT: LIBERT, FREDERICK

; TITLE OF INVENTION: ACTIVE AND INACTIVE CC-CHEMOKINES RECEPTOR

; TITLE OF INVENTION: AND NUCLEIC ACID MOLECULES ENCODING SAID RECEPTOR

; NUMBER OF SEQUENCES: 17

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Knobbe, Martens, Olson & Bear

; STREET: 620 Newport Center Drive 16th Floor

; CITY: Newport Beach

; STATE: CA

; COUNTRY: U.S.A.

; ZIP: 92660

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk

; COMPUTER: IBM PC compatible

; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: PatentIn Release #1.0, Version #1.25 (EPO)

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/08/833,752

; FILING DATE: 9-APR-1997

; CLASSIFICATION: 536

```
; ATTORNEY/AGENT INFORMATION:
; NAME: Altman, Daniel E
; REGISTRATION NUMBER: 34,115
; REFERENCE/DOCKET NUMBER:
; INFORMATION FOR SEQ ID NO: 5:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 352 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-833-752-5
```

```
Query Match          47.6%; Score 379; DB 4; Length 352;
Best Local Similarity 35.0%; Pred. No. 1.7e-34;
Matches 99; Conservative 7; Mismatches 27; Indels 150; Gaps 7;
```

```
QY      1 MDYQVSSPIYDINYYTSEPCQKINVKQIAA----- 30
          |||||||||||||||||||
Db      1 MDYQVSSPIYDINYYTSEPCQKINVKQIAARLLPLYSLVFIFGFVGNMLVILILINCKR 60

QY     31 -----YKCGLC-----AAAQWDFGNTMCQHQRVHGHHSYKCG--- 64
          |   |               |||||   | : | : : |
Db     61 LKSMTDIYLLNLAISDLFFLLTVPFWAHYAAAQWDFGNTMC--QLLTGLYFIGFFSGIFF 118

QY     65 -----LCTRSQKEGLHYTC 78
          : |||||||
Db    119 IILLTIDRYLAVVHAVFALKARTVTFGVVTSVITWVAVFASLPGIIFTRSQKEGLHYTC 178

QY     79 SSHFPYSQYQFWKNFQTLKI-----HQRVHGG----- 105
          |||||||||||||||   :: |
Db    179 SSHFPYSQYQFWKNFQTLKIVILGLVLP LLVMVICYS GILKTL LRCRNEKKRHRAVRLIF 238

QY    106 -----GGSYKCGLC---QEFFGLNNC SSSNRLDGHQRV 135
          |   |   |||||||   :|
Db    239 TIMIVYFLFWAPYNIVLLLNTFQEFFGLNNC SSSNRLDQAMQV 281
```

# RESULT 5

US-09-502-783A-2

; Sequence 2, Application US/09502783A

; Patent No. 6511826

; GENERAL INFORMATION:

; APPLICANT: Li, Yi

; APPLICANT: Ruben, Steven M.

; TITLE OF INVENTION: Polynucleotides. Encoding Human G-Protein Chemokine Receptor (CCR5)

; TITLE OF INVENTION: HDG NR10

; FILE REFERENCE: 1488.1150006

; CURRENT APPLICATION NUMBER: US/09/502,783A

; CURRENT FILING DATE: 2001-08-23

; PRIOR APPLICATION NUMBER: 08/466,343

; PRIOR FILING DATE: 1995-06-06

; NUMBER OF SEQ ID NOS: 9

; SOFTWARE: PatentIn version 3.0

; SEQ ID NO 2

; LENGTH: 352

; TYPE: PRT

; ORGANISM: Homo sapiens  
US-09-502-783A-2

Query Match 47.6%; Score 379; DB 4; Length 352;  
Best Local Similarity 35.0%; Pred. No. 1.7e-34;  
Matches 99; Conservative 7; Mismatches 27; Indels 150; Gaps 7;

```
Qy      1 MDYQVSSPIYDINYYTSEPCQKINVKQIAA----- 30
          ||||||||||||||||||||||||||||
Db      1 MDYQVSSPIYDINYYTSEPCQKINVKQIAARLLPPLYSLVFIFGFVGNMLVILILINCKR 60

Qy     31 -----YKCGLC-----AAAQWDFGNTMCQHQRVHGHHSYKCG--- 64
          |  |               ||||||||| | : | : : |
Db     61 LKSMTDIYLLNLAISDLFFLLTVPFWAHYAAAQWDFGNTMC--QLLTGLYFIGFFSGIFF 118

Qy     65 -----LCTRSQKEGLHYTC 78
          : |||||||||
Db    119 IILLTIDRYLAVVHAVFALKARTVTFGVVTSVITWVAVFASLPGIIFTRSQKEGLHYTC 178

Qy     79 SSHFPYSQYQFWKNFQTLKI-----HQRVHGG----- 105
          |||||||||||||||||
Db    179 SSHFPYSQYQFWKNFQTLKIVILGLVLPPLVMVICYSGILKTLRCRNEKKRHRAVRLIF 238

Qy    106 -----GGSYKCGLC-----QEFGNLCSSSNRLDGHQRV 135
          |  |               ||||||||| | : |
Db    239 TIMIVYFLWAPYNIVLLLNTFQEFGNLCSSSNRLDQAMQV 281
```

RESULT 6

US-09-796-202-1

; Sequence 1, Application US/09796202  
; Patent No. 6548636  
; GENERAL INFORMATION:  
; APPLICANT: Dragic, Tatjana  
; APPLICANT: Olson, William  
; TITLE OF INVENTION: SULFATED CCR5 PEPTIDES FOR HIV-1 INFECTION  
; FILE REFERENCE: 2048/61010/JPW/SHS  
; CURRENT APPLICATION NUMBER: US/09/796,202  
; CURRENT FILING DATE: 2001-02-28  
; NUMBER OF SEQ ID NOS: 17  
; SOFTWARE: PatentIn version 3.0  
; SEQ ID NO 1  
; LENGTH: 352  
; TYPE: PRT  
; ORGANISM: human  
US-09-796-202-1

Query Match 47.6%; Score 379; DB 4; Length 352;  
Best Local Similarity 35.0%; Pred. No. 1.7e-34;  
Matches 99; Conservative 7; Mismatches 27; Indels 150; Gaps 7;

```
Qy      1 MDYQVSSPIYDINYYTSEPCQKINVKQIAA----- 30
          ||||||||||||||||||||||||||||
Db      1 MDYQVSSPIYDINYYTSEPCQKINVKQIAARLLPPLYSLVFIFGFVGNMLVILILINCKR 60

Qy     31 -----YKCGLC-----AAAQWDFGNTMCQHQRVHGHHSYKCG--- 64
          |  |               ||||||||| | : | : : |
```

```

Db          61 LKSMTDIYLLNLAISDLFFLLTVPFWAHYAAAQWDFGNTMC--QLLTGLYFIGFFSGIFF 118
QY          65 -----LCTRSQKEGLHYTC 78
                                     : |||||
Db          119 IILLTIDRYLAVVHAVFALKARTVTFGVVTSVITWVVAVFASLPGIIFTRSQKEGLHYTC 178
QY          79 SSHFPYSQYQFWKNFQTLKI-----HQRVHGG----- 105
      |||||
Db          179 SSHFPYSQYQFWKNFQTLKIVILGLVLP LLVMVICYSGILKTLLRCRNEKKRHRAVRLIF 238
                                     :: |
QY          106 -----GGSYKCGLC----QEFFGLNNC SSSNRLDGHQRV 135
      | | |||||
Db          239 TIMIVYFLFWAPYNIVLLNTFQEFFGLNNC SSSNRLDQAMQV 281

```

RESULT 7

US-09-045-583-52

; Sequence 52, Application US/09045583

; Patent No. 6287805

; GENERAL INFORMATION:

; APPLICANT: Graham, Gerard J. et al.

; TITLE OF INVENTION: No. 6287805el Molecules of the G Protein-Coupled

; NUMBER OF SEQUENCES: 56

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: LAHIVE & COCKFIELD, LLP

; STREET: 28 State Street

; CITY: Boston

; STATE: Massachusetts

; COUNTRY: USA

; ZIP: 02109

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk

; COMPUTER: IBM PC compatible

; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: PatentIn Release #1.0, Version #1.25

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/09/045,583

; FILING DATE: 20-MAR-98

; CLASSIFICATION: 435

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER:

; FILING DATE:

; ATTORNEY/AGENT INFORMATION:

; NAME: Mandragouras, Amy E.

; REGISTRATION NUMBER: 36,207

; REFERENCE/DOCKET NUMBER: MNI-044

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: (617)227-7400

; TELEFAX: (617)742-4214

; INFORMATION FOR SEQ ID NO: 52:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 352 amino acids

; TYPE: amino acid

; TOPOLOGY: linear

; MOLECULE TYPE: peptide

; FRAGMENT TYPE: internal

US-09-045-583-52

Query Match 46.9%; Score 374; DB 3; Length 352;  
 Best Local Similarity 34.6%; Pred. No. 6e-34;  
 Matches 98; Conservative 8; Mismatches 27; Indels 150; Gaps 7;

```

Qy      1 MDYQVSSPIYDINYYTSEPCQKINVKQIAA----- 30
          |||||:|||||
Db      1 MDYQVSSPIYDIDYYTSEPCQKINVKQIAARLLPPLYSLVFIFGFVGNMLVILILINCKR 60

Qy     31 -----YKCGLC-----AAAQWDFGNTMCQHQRVHGHHSYKCG--- 64
          |  |               |||||  | : | : |
Db     61 LKSMTDIYLLNLAISDLFFLLTVPFWAHYAAAQWDFGNTMC--QLLTGLYFIGFFSGIFF 118

Qy     65 -----LCTRSQKEGLHYTC 78
          : |||||
Db    119 IILLTIDRYLAIVHAVFALKARTVTFGVVTSVITWVAVFASLPGIIFTRSQKEGLHYTC 178

Qy     79 SSHFPYSQYQFWKNFQTLKI-----HQRVHGG----- 105
          |||||               :: |
Db    179 SSHFPYSQYQFWKNFQTLKIVILGLVLP LLVMVICYSGILKTLRLCRNEKKRHRAVRLIF 238

Qy    106 -----GGSYKCGLC----QEFGGLNNCSSSNRLDGHQRV 135
          |  |               |||||  :|
Db    239 TIMIVYFLFWAPYNIVLLLNTFQEFGGLNNCSSSNRLDQAMQV 281
  
```

RESULT 8

US-09-534-185-52

; Sequence 52, Application US/09534185

; Patent No. 6403767

; GENERAL INFORMATION:

; APPLICANT: Graham, Gerard J. et al.

; TITLE OF INVENTION: No. 6403767el Molecules of the G Protein-Coupled  
 ; Heptahelical Receptor Superfamily and Uses  
 ; Therefor

; NUMBER OF SEQUENCES: 56

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: LAHIVE & COCKFIELD, LLP

; STREET: 28 State Street

; CITY: Boston

; STATE: Massachusetts

; COUNTRY: USA

; ZIP: 02109

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk

; COMPUTER: IBM PC compatible

; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: PatentIn Release #1.0, Version #1.25

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/09/534,185

; FILING DATE: 24-Mar-2000

; CLASSIFICATION: <Unknown>

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: 09/045,583

; FILING DATE: <Unknown>

; ATTORNEY/AGENT INFORMATION:

; NAME: Mandragouras, Amy E.



```

;      REGISTRATION NUMBER: 36,207
;      REFERENCE/DOCKET NUMBER: MNI-044
;      TELECOMMUNICATION INFORMATION:
;      TELEPHONE: (617)227-7400
;      TELEFAX: (617)742-4214
;      INFORMATION FOR SEQ ID NO: 52:
;      SEQUENCE CHARACTERISTICS:
;      LENGTH: 352 amino acids
;      TYPE: amino acid
;      TOPOLOGY: linear
;      MOLECULE TYPE: peptide
;      FRAGMENT TYPE: internal
;      SEQUENCE DESCRIPTION: SEQ ID NO: 52:
US-09-534-185-52

```

```

Query Match          46.9%;  Score 374;  DB 4;  Length 352;
Best Local Similarity 34.6%;  Pred. No. 6e-34;
Matches 98;  Conservative 8;  Mismatches 27;  Indels 150;  Gaps 7;

```

```

Qy      1 MDYQVSSPIYDINYYTSEPCQKINVKQIAA----- 30
          |||
Db      1 MDYQVSSPIYDIDYYTSEPCQKINVKQIAARLLPPLYSLVFIFGFVGNMLVILILINCKR 60

Qy      31 -----YKCGLC-----AAAQWDFGNTMCQHQRVHGHHHSYKCG--- 64
          | |
Db      61 LKSMTDIYLLNLAISDLFFLLTVPFWAHYAAAQWDFGNTMC--QLLTGLYFIGFFSGIFF 118

Qy      65 -----LCTRSQKEGLHYTC 78
          : |||
Db      119 IILLTIDRYLAIVHAVFALKARTVTFGVVTSVITWVVAVFASLPGIIFTRSQKEGLHYTC 178

Qy      79 SSHFPYSQYQFWKNFQTLKI-----HQRVHGG----- 105
          |||
Db      179 SSHFPYSQYQFWKNFQTLKIVILGLVLP LLVMVICYS GILKTL LRCRNEKKRHRVRLIF 238

Qy      106 -----GGSYKCGLC----QEFG LNNC SSSNRLDGHQRV 135
          | |
Db      239 TIMIVYFLFWAPYNIVLLLNTFQEFG LNNC SSSNRLDQAMQV 281

```

RESULT 9

US-08-466-343D-2

; Sequence 2, Application US/08466343D

; Patent No. 6025154

; GENERAL INFORMATION:

; APPLICANT: LI, Yi

; TITLE OF INVENTION: POLYNUCLEOTIDES ENCODING HUMAN G-PROTEIN

; TITLE OF INVENTION: CHEMOKINE RECEPTOR HDG NR10 (AS AMENDED)

; NUMBER OF SEQUENCES: 9

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: STERNE, KESSLER, GOLDSTEIN & FOX P.L.L.C.

; STREET: 1100 NEW YORK AVE., NW, SUITE 600

; CITY: WASHINGTON

; STATE: DC

; COUNTRY: USA

; ZIP: 20005

; COMPUTER READABLE FORM:

```

; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/466,343D
; FILING DATE: 06-JUN-1995
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: STEFFE, ERIC K.
; REGISTRATION NUMBER: 36,688
; REFERENCE/DOCKET NUMBER: 1488.1150000/EKS/KLM
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202) 371-2600
; TELEFAX: (202) 371-2540
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 352 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-466-343D-2

```

```

Query Match          46.8%; Score 373; DB 3; Length 352;
Best Local Similarity 34.6%; Pred. No. 7.8e-34;
Matches 98; Conservative 7; Mismatches 28; Indels 150; Gaps 7;

```

```

Qy      1 MDYQVSSPIYDINYYTSEPCQKINVKQIAA----- 30
          |||
Db      1 MDYQVSSPIYDINYYTSEPCPKINVKQIAARLLPPLYSLVFIFGFVGNMLVILILINCQR 60
          |||

Qy      31 -----YKCGLC-----AAAQWDFGNTMCQHQRVHGHHSYKCG--- 64
          | | |||
Db      61 LESMTDIYLLNLAISDLFFLLTVPFWAHYAAAQWDFGNTMC--QLLTGLYFIGFFSGIFF 118
          | | |||

Qy      65 -----LCTRSQKEGLHYTC 78
          : |||
Db      119 IILLTIDRYLAIVHAVFALKARTVTFGVVTSVITWVVAVFASLPGIIFTRSQKEGLHYTC 178
          : |||

Qy      79 SSHFPYSQYQFWKNFQTLKI-----HQRVHGG----- 105
          |||
Db      179 SSHFPYSQYQFWKNFQTLKIVILGLVLP LLVMVICYS GILKTLLRCRNEKKRHRVRLIF 238
          : |

Qy      106 -----GGSYKCGLC---QEFFGLNNC SSSNRLDGHQRV 135
          | | |||
Db      239 TIMIVYFLFWAPYNIVLLLNTFQEFFGLNNC SSSNRLDQAMQV 281
          : |

```

# RESULT 10

US-09-517-605-5

```

; Sequence 5, Application US/09517605
; Patent No. 6391567
; GENERAL INFORMATION:
; APPLICANT: Littman, Dan R.
; APPLICANT: Kwon, Douglas S.
; APPLICANT: van Kooyk, Yvette
; APPLICANT: Geijtenbeck, Theo

```

```

; TITLE OF INVENTION: METHODS OF USING A FACILITATOR OF RETROVIRAL ENTRY INTO
; TITLE OF INVENTION: CELLS
; FILE REFERENCE: 1049-1-017
; CURRENT APPLICATION NUMBER: US/09/517,605
; CURRENT FILING DATE: 2000-03-02
; NUMBER OF SEQ ID NOS: 17
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 5
; LENGTH: 352
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-517-605-5

```

```

Query Match          45.0%; Score 359; DB 4; Length 352;
Best Local Similarity 33.9%; Pred. No. 2.9e-32;
Matches 96; Conservative 8; Mismatches 29; Indels 150; Gaps 7;

Qy      1 MDYQVSSPIYDINYYTSEPCQKINVKQIAA----- 30
        ||||| |||:| |||||
Db      1 MDYQVSSPTYDIDYDTSEPCQKINVKQIAARLLPLYSLVFIFGFVGNMLVILVLINCKR 60

Qy      31 -----YKCGLC-----AAAQWDFGNTMCQHQRVHGHHSYKCG--- 64
        | | ||||| | : | : |
Db      61 LKSMTDIYLLNLAISDLFFLLTVPFWAHYAAAQWDFGNTMC--QLLTGLYFIGFFSGIFF 118

Qy      65 -----LCTRSQKEGLHYTC 78
        : |||||
Db      119 IILLTIDRYLAIVHAVFALKARTVTFGVVTSVITWVAVFASLPGIIFTRSQKEGLHYTC 178

Qy      79 SSHFPYSQYQFWKNFQTLKI-----HQRVHGG----- 105
        ||||| ||||| : |
Db      179 SSHFPYSQYQFWKNFQTLKIVILGLVLP LLVMVICYSGILKTLLRCRNEKKRHRVRLIF 238

Qy      106 -----GGSYKCGLC---QEFGFLNNCSSSNRLDGHQRV 135
        | | ||||| : |
Db      239 TIMIVYFLFWAPYNIVLLLNTFQEFGFLNNCSSSNRLDQAMQV 281

```

RESULT 11

US-08-724-984A-2

; Sequence 2, Application US/08724984A

; Patent No. 6388055

; GENERAL INFORMATION:

; APPLICANT: Derk Bergsma, Mary Brawner, and Usman Shabon

; TITLE OF INVENTION: No. 6388055el Mouse Genomic Clone of the CC-

; TITLE OF INVENTION: CKR5 Receptor

; NUMBER OF SEQUENCES: 5

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: SmithKline Beecham Corporation

; STREET: 709 Swedeland Road, P.O. Box 1539

; CITY: King of Prussia

; STATE: PA

; COUNTRY: USA

; ZIP: 19406-0939

; COMPUTER READABLE FORM:

; MEDIUM TYPE: DISKETTE, 3.5 INCH, 1.44 Mb STORAGE

; COMPUTER: IBM 486

```

;   OPERATING SYSTEM:  WINDOWS FOR WORKGROUPS
;   SOFTWARE:  MICROSOFT WORD
;   CURRENT APPLICATION DATA:
;   APPLICATION NUMBER:  US/08/724,984A
;   FILING DATE:  October 3, 1996
;   CLASSIFICATION:  800
;   PRIOR APPLICATION DATA:
;   APPLICATION NUMBER:
;   FILING DATE:
;   ATTORNEY/AGENT INFORMATION:
;   NAME:  William T. Han
;   REGISTRATION NUMBER:  34,344
;   REFERENCE/DOCKET NUMBER:  ATG50023
;   TELECOMMUNICATION INFORMATION:
;   TELEPHONE:  610 270 5024
;   TELEFAX:  610 270 5090
;   INFORMATION FOR SEQ ID NO:  2:
;   SEQUENCE CHARACTERISTICS:
;   LENGTH:  354
;   TYPE:  Amino Acid
;   TOPOLOGY:  Linear
US-08-724-984A-2

```

```

Query Match          34.1%;  Score 272;  DB 4;  Length 354;
Best Local Similarity 27.9%;  Pred. No. 1.6e-22;
Matches  80;  Conservative  12;  Mismatches  29;  Indels  166;  Gaps  8;

```

```

Qy      1 MDYQVSSP--IYDINYTSEPCQKINVKQIAA----- 30
      ||:| | | |||:| | |||||
Db      1 MDFQGSVPYTYIDIDYGMSAPCQKINVKQIAAQLLPPLYSLVFIFGFAGNMMVFLILISC 60

Qy     31 -----YKCGLC-----AAAQWDFGNTMCQHQRVHGHHSYKCG- 64
      | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db     61 KKLKSVTDIYLLNLAISDLLFLLTLFPWAHYAANEWIFGNIMC--KVFTGVYHIGYFGGI 118

Qy     65 -----LCTRSQKEGLHY 76
      : ||||| |
Db    119 FFIILLTIDRYLAIVHAVFALKVTTVNFVITSVVTWVAVFASLPEIIFTRSQKEGFHY 178

Qy     77 TCSSHFPYSQYQFWKNFQTLK-----IHQRVHGGGGSYKC----- 111
      ||| |||:| | |||:| | | | | | | | | | | | | | | | | | | | | |
Db    179 TCSPHFPHTQYHFWSFQTLKMLVLSLILPLLVMIIICYSGLHT-----LFRCRNEKK 231

Qy    112 -----GLCQEFFGLNNCSSSNRLD 130
      ||||| |
Db    232 RHRAVRLIFAIMIVYFLFWTPYNIVLLLTTFQEFFGLNNCSSSNRLD 278

```

# RESULT 12

US-08-833-752-4

```

; Sequence 4, Application US/08833752
; Patent No. 6448375
; GENERAL INFORMATION:
;   APPLICANT:  SAMSON, MICHEL
;   APPLICANT:  PARMENTIER, MARC
;   APPLICANT:  VASSART, GILBERT
;   APPLICANT:  LIBERT, FREDERICK

```

```

; TITLE OF INVENTION: ACTIVE AND INACTIVE CC-CHEMOKINES RECEPTOR
; TITLE OF INVENTION: AND NUCLEIC ACID MOLECULES ENCODING SAID RECEPTOR
; NUMBER OF SEQUENCES: 17
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Knobbe, Martens, Olson & Bear
; STREET: 620 Newport Center Drive 16th Floor
; CITY: Newport Beach
; STATE: CA
; COUNTRY: U.S.A.
; ZIP: 92660
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25 (EPO)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/833,752
; FILING DATE: 9-APR-1997
; CLASSIFICATION: 536
; ATTORNEY/AGENT INFORMATION:
; NAME: Altman, Daniel E
; REGISTRATION NUMBER: 34,115
; REFERENCE/DOCKET NUMBER:
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 184 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-833-752-4

```

```

Query Match          32.4%; Score 258; DB 4; Length 184;
Best Local Similarity 34.9%; Pred. No. 2.7e-21;
Matches 65; Conservative 4; Mismatches 13; Indels 104; Gaps 4;

```

```

Qy      1 MDYQVSSPIYDINYYTSEPCQKINVQIAA----- 30
          |||||||||||||||||||
Db      1 MDYQVSSPIYDINYYTSEPCQKINVQIAARLLPPLYSLVFIFGFVGNMLVILILINCKR 60

Qy      31 -----YKCGLC-----AAAQWDFGNTMCQHQRVHGHHSYKCG--- 64
          | |                      ||||| | : | : |
Db      61 LKSMTDIYLLNLAISDLFFLLTVPFWAHYAAAQWDFGNTMC--QLLTGLYFIGFFSGIFF 118

Qy      65 -----LCTRSQKEGLHYTC 78
          : |||||
Db      119 IILLTIDRYLAVVHAVFALKARTVTFGVVTSVITWVAVFASLPGIIFTRSQKEGLHYTC 178

Qy      79 SSHFPY 84
          |||||
Db      179 SSHFPY 184

```

```

RESULT 13
US-09-087-232A-17
; Sequence 17, Application US/09087232A
; Patent No. 6153431
; GENERAL INFORMATION:

```

```

;   APPLICANT:  Quillent et al.
;   TITLE OF INVENTION:  HUMAN IMMUNODEFICIENCY VIRUS CO-RECEPTOR
;   TITLE OF INVENTION:  VARIANTS ASSOCIATED WITH RESISTANCE TO VIRUS
INFECTION.
;   NUMBER OF SEQUENCES:  23
;   CORRESPONDENCE ADDRESS:
;       ADDRESSEE:  Baker & Botts, L.L.P. attn. Lisa Kole
;       STREET:  30 Rockefeller Plaza
;       CITY:  New York
;       STATE:  New York
;       COUNTRY:  USA
;       ZIP:  10112
;   COMPUTER READABLE FORM:
;       MEDIUM TYPE:  Floppy disk
;       COMPUTER:  IBM PC compatible
;       OPERATING SYSTEM:  PC-DOS/MS-DOS
;       SOFTWARE:  PatentIn Release #1.0, Version #1.30 (EPO)
;   CURRENT APPLICATION DATA:
;       APPLICATION NUMBER:  US/09/087,232A
;       FILING DATE:  28 MAY 1998
;       CLASSIFICATION:  435
;   PRIOR APPLICATION DATA:
;       APPLICATION NUMBER:  60/048,057
;       FILING DATE:  30 MAY 1997
;   ATTORNEY/AGENT INFORMATION:
;       NAME:  KOLE, LISA B.
;       REGISTRATION NUMBER:  35,225
;       REFERENCE/DOCKET NUMBER:  AP 31115
;   TELECOMMUNICATION INFORMATION:
;       TELEPHONE:  (212) 408-2628
;       TELEFAX:  (212) 765-2519
;   INFORMATION FOR SEQ ID NO:  17:
;       SEQUENCE CHARACTERISTICS:
;           LENGTH: 215 amino acids
;           TYPE:  amino acid
;           TOPOLOGY:  linear
;       MOLECULE TYPE:  protein
US-09-087-232A-17

```

```

Query Match          32.4%;  Score 258;  DB 3;  Length 215;
Best Local Similarity 34.9%;  Pred. No. 3.3e-21;
Matches 65;  Conservative 4;  Mismatches 13;  Indels 104;  Gaps 4;

```

```

QY      1 MDYQVSSPIYDINYYTSEPCQKINVKQIAA----- 30
          |||
Db      1 MDYQVSSPIYDINYYTSEPCQKINVKQIAARLLPLYSLVFIFGFVGNMLVILILINCKR 60

QY     31 -----YKCGLC-----AAAQWDFGNTMCQHQRVHGHHSYKCG--- 64
          |  |
Db     61 LKSMTDIYLLNLAISDLFFLLTVPFWAHYAAAQWDFGNTMC--QLLTGLYFIGFFSGIFF 118

QY     65 -----LCTRSQKEGLHYTC 78
          : |||
Db    119 IILLTIDRYLAVVHAVFALKARTVTFGVVTSVITWVVAVFASLPGIIFTRSQKEGLHYTC 178

QY     79 SSHFPY 84
          |||||

```

Db 179 SSHFPY 184

RESULT 14  
US-08-833-752-6  
; Sequence 6, Application US/08833752  
; Patent No. 6448375  
; GENERAL INFORMATION:  
; APPLICANT: SAMSON, MICHEL  
; APPLICANT: PARMENTIER, MARC  
; APPLICANT: VASSART, GILBERT  
; APPLICANT: LIBERT, FREDERICK  
; TITLE OF INVENTION: ACTIVE AND INACTIVE CC-CHEMOKINES RECEPTOR  
; TITLE OF INVENTION: AND NUCLEIC ACID MOLECULES ENCODING SAID RECEPTOR  
; NUMBER OF SEQUENCES: 17  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Knobbe, Martens, Olson & Bear  
; STREET: 620 Newport Center Drive 16th Floor  
; CITY: Newport Beach  
; STATE: CA  
; COUNTRY: U.S.A.  
; ZIP: 92660  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.25 (EPO)  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/833,752  
; FILING DATE: 9-APR-1997  
; CLASSIFICATION: 536  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Altman, Daniel E  
; REGISTRATION NUMBER: 34,115  
; REFERENCE/DOCKET NUMBER:  
; INFORMATION FOR SEQ ID NO: 6:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 215 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
US-08-833-752-6

Query Match 32.4%; Score 258; DB 4; Length 215;  
Best Local Similarity 34.9%; Pred. No. 3.3e-21;  
Matches 65; Conservative 4; Mismatches 13; Indels 104; Gaps 4;

```
QY      1 MDYQVSSPIYDINYYTSEPCQKINVKQIAA----- 30
          |||
Db      1 MDYQVSSPIYDINYYTSEPCQKINVKQIAARLLPPLYSLVFIFGFVGNMLVILILINCKR 60
QY     31 -----YKCGLC-----AAAQWDFGNTMCQHQRVHGHGHHHSYKCG--- 64
          |
Db     61 LKSMTDIYLLNLAISDLFFLLTVPFWAHYAAAQWDFGNTMC--QLLTGLYFIGFFSGIFF 118
QY     65 -----LCTRSQKEGLHYTC 78
          : |||
```

Db 119 IILLTIDRYLAVVHAVFALKARTVTFGVVTSVITWVVAVFASLPGIIFTRSQKEGLHYTC 178  
Qy 79 SSHFPY 84  
|||||  
Db 179 SSHFPY 184

RESULT 15

US-09-087-232A-15  
; Sequence 15, Application US/09087232A  
; Patent No. 6153431  
; GENERAL INFORMATION:  
; APPLICANT: Quillent et al.  
; TITLE OF INVENTION: HUMAN IMMUNODEFICIENCY VIRUS CO-RECEPTOR  
; TITLE OF INVENTION: VARIANTS ASSOCIATED WITH RESISTANCE TO VIRUS  
INFECTION.  
; NUMBER OF SEQUENCES: 23  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Baker & Botts, L.L.P. attn. Lisa Kole  
; STREET: 30 Rockefeller Plaza  
; CITY: New York  
; STATE: New York  
; COUNTRY: USA  
; ZIP: 10112  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.30 (EPO)  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/09/087,232A  
; FILING DATE: 28 MAY 1998  
; CLASSIFICATION: 435  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 60/048,057  
; FILING DATE: 30 MAY 1997  
; ATTORNEY/AGENT INFORMATION:  
; NAME: KOLE, LISA B.  
; REGISTRATION NUMBER: 35,225  
; REFERENCE/DOCKET NUMBER: AP 31115  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (212) 408-2628  
; TELEFAX: (212) 765-2519  
; INFORMATION FOR SEQ ID NO: 15:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 100 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
US-09-087-232A-15

Query Match 22.9%; Score 182.5; DB 3; Length 100;  
Best Local Similarity 43.0%; Pred. No. 3.8e-13;  
Matches 43; Conservative 0; Mismatches 4; Indels 53; Gaps 2;

Qy 1 MDYQVSSPIYDINYTSEPCQKINVKQIAA----- 30  
|||||



Db           1 MDYQVSSPIYDINYYTSEPCQKINVKQIAARLLPPLYSLVFIFGFVGNMLVILILINCKR 60  
Qy           31 -----YKCGLC-----AAAQWDFGNTM 47  
              |   |                               | | | | | | | | | |  
Db           61 LKSMTDIYLLNLAISDLFFLLTVPFWAHYAAAQWDFGNTM 100

Search completed: September 17, 2004, 16:49:03  
Job time : 23.5186 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: September 17, 2004, 16:42:06 ; Search time 22.4542 Seconds  
(without alignments)  
591.177 Million cell updates/sec

Title: US-10-057-890A-10  
Perfect score: 797  
Sequence: 1 MDYQVSSPIYDINYYTSEPC.....GLNNCSSSNRLDGHQRVHAA 138

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 283366 seqs, 96191526 residues

Total number of hits satisfying chosen parameters: 283366

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : PIR\_78:\*  
1: pir1:\*  
2: pir2:\*  
3: pir3:\*  
4: pir4:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query		DB	ID	Description
		Match	Length			
1	379	47.6	352	2	A43113	chemokine (C-C) re
2	154	19.3	201	2	I57505	zinc finger protei
3	144	18.1	693	2	I37570	zinc finger protei
4	143.5	18.0	319	2	T46469	hypothetical prote
5	141	17.7	555	2	I53869	zinc finger protei
6	138.5	17.4	223	2	F31201	GLI-related finger
7	136.5	17.1	348	2	I38599	zinc finger protei
8	135.5	17.0	572	2	I39311	Kruppel-type zinc
9	134	16.8	710	2	I48668	zinc finger protei
10	133.5	16.8	475	2	S03679	finger protein (cl
11	133.5	16.8	636	2	I48689	gene NK10 protein
12	133.5	16.8	701	2	T14757	hypothetical prote
13	131.5	16.5	686	2	A34612	zinc finger protei

14	131	16.4	261	2	S70006	finger protein zfo
15	129.5	16.2	614	2	JH0500	zinc finger protei
16	128.5	16.1	386	2	T12527	hypothetical prote
17	128.5	16.1	580	2	A37107	spermatogenesis pr
18	127.5	16.0	469	2	I38600	zinc finger protei
19	127.5	16.0	1173	2	I50620	prockr2 - chicken
20	127	15.9	728	2	A48830	probable transcrip
21	126.5	15.9	732	2	S47073	finger protein HZF
22	124	15.6	367	2	S06582	finger protein (cl
23	124	15.6	589	2	I38598	zinc finger protei
24	123.5	15.5	209	2	S47068	finger protein HZF
25	123.5	15.5	347	2	S00549	developmental cont
26	123.5	15.5	378	2	S33994	finger protein ZNF
27	122.5	15.4	488	2	S47072	finger protein HZF
28	122.5	15.4	594	2	T12488	hypothetical prote
29	122.5	15.4	803	2	S26823	zinc finger protei
30	122	15.3	1350	2	S00647	finger protein - A
31	120.5	15.1	169	2	A39240	finger protein mfg
32	120.5	15.1	229	2	A48927	Kruppel-like zinc
33	120.5	15.1	399	2	S47071	finger protein HZF
34	120.5	15.1	546	2	I49636	DNA-binding protei
35	120.5	15.1	1191	2	S35305	zinc finger protei
36	120	15.1	435	2	S00833	finger protein (cl
37	120	15.1	654	2	A57785	finger protein ZNF
38	119.5	15.0	194	2	I53859	zinc finger protei
39	119.5	15.0	337	2	S60520	finger protein ZNF
40	119.5	15.0	393	2	JN0533	finger protein pML
41	119.5	15.0	576	2	A48157	renal transcriptio
42	119.5	15.0	671	2	JE0288	krueppel-type zinc
43	118.5	14.9	107	2	I39315	zinc-finger protei
44	118.5	14.9	120	2	G02493	KR-ZNF1 - human (f
45	118.5	14.9	183	2	S70007	finger protein zfo

## ALIGNMENTS

### RESULT 1

A43113

chemokine (C-C) receptor 5 - human

N;Alternate names: C-C CKR-5; CCR5

C;Species: Homo sapiens (man)

C;Date: 12-Jul-1996 #sequence\_revision 12-Jul-1996 #text\_change 20-Jun-2000

C;Accession: A43113; S71808; A58834; A58832; G02653; A58833

R;Samson, M.; Labbe, O.; Mollereau, C.; Vassart, G.; Parmentier, M.

Biochemistry 35, 3362-3367, 1996

A;Title: Molecular cloning and functional expression of a new human CC-chemokine receptor gene.

A;Reference number: A43113; MUID:96241590; PMID:8639485

A;Accession: A43113

A;Molecule type: mRNA

A;Residues: 1-352 <SAM1>

A;Cross-references: GB:X91492; NID:g1262810; PIDN:CAA62796.1; PID:g1262811

R;Samson, M.; Libert, F.; Doranz, B.J.; Rucker, J.; Liesnard, C.; Farber, C.M.; Saragosti, S.; Lapoumeroulie, C.; Cognaux, J.; Forceille, C.; Muyldermans, G.; Verhofstede, C.; Burtonboy, G.; Georges, M.; Imai, T.; Rana, S.; Yi, Y.; Smyth, R.J.; Collman, R.G.; Doms, R.W.; Vassart, G.; Parmentier, M.

Nature 382, 722-725, 1996

A;Title: Resistance to HIV-1 infection in caucasian individuals bearing mutant alleles of the CCR-5 chemokine receptor gene.

A;Reference number: S71808; MUID:96345670; PMID:8751444

A;Accession: S71808

A;Status: nucleic acid sequence not shown; not compared with conceptual translation

A;Molecule type: DNA

A;Residues: 182-206;207-230 <SAM2>

A;Accession: A58834

A;Status: nucleic acid sequence not shown; not compared with conceptual translation

A;Molecule type: DNA

A;Residues: 1-184,'IKDSHLGAGPAAACHGHLLLGPNKNSASVSK' <SAM3>

A;Cross-references: GB:X99393; NID:g1524062; PIDN:CAA67767.1; PID:g1524063

A;Note: this frameshift mutation results in a non-functional receptor but confers a degree of resistance to HIV-1 infection; it has an allele frequency of 0.09 or more in some caucasian populations and may have had a selective advantage by conferring resistance to Yersinia plague infections

R;Combadiere, C.; Ahuja, S.K.; Tiffany, H.L.; Murphy, P.M.

J. Leukoc. Biol. 60, 147-152, 1996

A;Title: Cloning and functional expression of CC CKR5, a human monocyte CC chemokine receptor selective for MIP-1alpha, MIP-1beta, and RANTES.

A;Reference number: A58832; MUID:96295970; PMID:8699119

A;Accession: A58832

A;Molecule type: mRNA

A;Residues: 1-352 <COM1>

A;Cross-references: GB:U57840; NID:g1502408; PIDN:AAB17071.1; PID:g1502409

A;Experimental source: clone 8, endotoxin-stimulated peripheral blood monocytes

R;Combadiere, C.

submitted to the EMBL Data Library, May 1996

A;Reference number: H01541

A;Accession: G02653

A;Status: translated from GB/EMBL/DDBJ

A;Molecule type: mRNA

A;Residues: 1-89,'L',91-352 <COM2>

A;Cross-references: EMBL:U57840

R;Raport, C.J.; Gosling, J.; Schweickart, V.L.; Gray, P.W.; Charo, I.F.

J. Biol. Chem. 271, 17161-17166, 1996

A;Title: Molecular cloning and functional characterization of a novel human CC chemokine receptor (CCR5) for RANTES, MIP-1beta, and MIP-1alpha.

A;Reference number: A58833; MUID:96291862; PMID:8663314

A;Accession: A58833

A;Molecule type: mRNA

A;Residues: 1-352 <RAP>

A;Cross-references: GB:U54994; NID:g1457945; PIDN:AAC50598.1; PID:g1457946

C;Comment: This is a receptor for chemokines MIP-1alpha (see PIR:A30574), MIP-1beta (see PIR:A31767), and RANTES (see PIR:A28815).

C;Comment: Macrophage- and dual-tropic strains of HIV-1 bind to a complex of chemokine (C-C) receptor 5 and T-cell surface glycoprotein CD4 (see PIR:RWHUT4).

C;Genetics:

A;Gene: GDB:CMKBR5; CCR5; CKR-5; CC-CKR-5; CKR5; ChemR13

A;Cross-references: GDB:1230510; OMIM:601373

A;Map position: 3p21-3p21

C;Function:

A;Description: G protein-coupled receptor for chemokines MIP-1alpha, MIP-1beta and RANTES

A;Note: probably acts to control granulocyte proliferation and differentiation  
 C;Superfamily: vertebrate rhodopsin  
 C;Keywords: AIDS; G protein-coupled receptor; glycoprotein; phosphoprotein;  
 transmembrane protein  
 F;32-56/Domain: transmembrane #status predicted <TM1>  
 F;67-87/Domain: transmembrane #status predicted <TM2>  
 F;103-124/Domain: transmembrane #status predicted <TM3>  
 F;142-166/Domain: transmembrane #status predicted <TM4>  
 F;193-218/Domain: transmembrane #status predicted <TM5>  
 F;236-257/Domain: transmembrane #status predicted <TM6>  
 F;285-300/Domain: transmembrane #status predicted <TM7>  
 F;20-269,101-178/Disulfide bonds: #status predicted  
 F;268/Binding site: carbohydrate (Asn) (covalent) #status predicted  
 F;336,337,342/Binding site: phosphate (Ser) (covalent) #status predicted  
 F;340,343/Binding site: phosphate (Thr) (covalent) #status predicted

Query Match 47.6%; Score 379; DB 2; Length 352;  
 Best Local Similarity 35.0%; Pred. No. 2.5e-29;  
 Matches 99; Conservative 7; Mismatches 27; Indels 150; Gaps 7;

```

Qy      1 MDYQVSSPIYDINYYTSEPCQKINVKQIAA----- 30
          |||
Db      1 MDYQVSSPIYDINYYTSEPCQKINVKQIAARLLPPLYSLVFIFGFVGNMLVILILINCKR 60

Qy      31 -----YKCGLC-----AAAQWDFGNTMCQHQRVHGHHHSYKCG--- 64
          |  |
Db      61 LKSMTDIYLLNLAISDLFFLLTVPFWAHYAAAQWDFGNTMC--QLLTGLYFIGFFSGIFF 118

Qy      65 -----LCTRSQKEGLHYTC 78
          : |||
Db      119 IILLTIDRYLAVVHAVFALKARTVTFGVVTSVITWVAVFASLPGLIFTRSQKEGLHYTC 178

Qy      79 SSHFPYSQYQFWKNFQTLKI-----HQRVHGG----- 105
          |||
Db      179 SSHFPYSQYQFWKNFQTLKIVILGLVLP LLVMVICYSGILKTLRLCRNEKKRHRVRLIF 238

Qy      106 -----GGSYKCGLC-----QEFFGLNNC SSSNRLDGHQRV 135
          |  |
Db      239 TIMIVYFLFWAPYNIVLLNTFQEFFGLNNC SSSNRLDQAMQV 281
  
```

## RESULT 2

I57505

zinc finger protein - mouse

C;Species: Mus musculus (house mouse)

C;Date: 02-Aug-1996 #sequence\_revision 02-Aug-1996 #text\_change 05-Nov-1999

C;Accession: I57505

R;Ernoul-Lange, M.; Kress, M.; Hamer, D.

Mol. Cell. Biol. 10, 418-421, 1990

A;Title: A gene that encodes a protein consisting solely of zinc finger domains is preferentially expressed in transformed mouse cells.

A;Reference number: I57505; MUID:90097859; PMID:2104662

A;Accession: I57505

A;Status: preliminary; translated from GB/EMBL/DDBJ

A;Molecule type: mRNA

A;Residues: 1-201 <RES>

A;Cross-references: GB:M32057; NID:g199766; PIDN:AAA39729.1; PID:g199767

C;Genetics:  
A;Gene: MOK2

Query Match 19.3%; Score 154; DB 2; Length 201;  
Best Local Similarity 37.0%; Pred. No. 9.6e-08;  
Matches 44; Conservative 13; Mismatches 44; Indels 18; Gaps 6;

```
Qy      26 KQIAAYKCGLCAAAQWDFGNTMCQHQRVHGHHSYKCGLCCTR--SQKEGLHYTCSSHF- 82
      | | :||: : : : : |||| | || | | : |
Db      60 KPYACEECGMSFSQR----SNLHIHQRVH-TGERPYKCGECGKGFSSQSSNLHIHRCHTTG 114

Qy      83 --PYSQYQFWKNF---QTLKIHQRVHGGGGSYKCGLCQEFFGLNNCSSSNRLDGHQRVH 136
      || | : | | | :|| || | | || | : | | :| ||||
Db     115 EKPYQCYECGKGFSSQSSDLRIHLRVHTGEKPYHCCKGQGF-----SQSSKLLIHQRVH 168
```

#### RESULT 3

I37570

zinc finger protein - human (fragment)

C;Species: Homo sapiens (man)

C;Date: 29-May-1998 #sequence\_revision 29-May-1998 #text\_change 01-Dec-2000

C;Accession: I37570

R;Abrink, M.; Aveskogh, M.; Hellman, L.

DNA Cell Biol. 14, 125-136, 1995

A;Title: Isolation of cDNA clones for 42 different Kruppel-related zinc finger proteins expressed in the human monoblast cell line U-937.

A;Reference number: I37566; MUID:95169271; PMID:7865130

A;Accession: I37570

A;Status: preliminary; translated from GB/EMBL/DDBJ

A;Molecule type: mRNA

A;Residues: 1-693 <RES>

A;Cross-references: EMBL:X78927; NID:g498726; PIDN:CAA55527.1; PID:g498727

C;Genetics:

A;Gene: HZF4

C;Superfamily: zinc finger protein ZFP-36; LIM metal-binding repeat homology

Query Match 18.1%; Score 144; DB 2; Length 693;  
Best Local Similarity 33.3%; Pred. No. 3e-06;  
Matches 39; Conservative 15; Mismatches 33; Indels 30; Gaps 7;

```
Qy      21 QKINVKQIAAYKCGLCAAA-QWDFGNTMCQHQRVHGHHSYKCGLCSTRSQKEGLHYTCS 79
      ||:: : |||| | :| | |||| | || | | ||::
Db     554 QKVHTGE-KPYKCGECGKGFKWSLNLDL--HQRVH-TGEKPYTCGAC-----GKHFSQA 603

Qy      80 SHFPYSQYQFWKNFQTLKIHQRVHGGGGSYKCGLCQEFFGLNNCSSSNRLDGHQRVH 136
      | :||: || | || | :| : | | :| | :||
Db     604 S-----SLQLHQSVHTGEKPYKCDVCGKVF-----SRSSQLQYHRRVH 641
```

#### RESULT 4

T46469

hypothetical protein DKFZp434G1930.1 - human (fragment)

C;Species: Homo sapiens (man)

C;Date: 04-Feb-2000 #sequence\_revision 04-Feb-2000 #text\_change 04-Feb-2000

C;Accession: T46469

R;Blum, H.; Bauersachs, S.; Mewes, H.W.; Gassenhuber, J.; Wiemann, S.

submitted to the Protein Sequence Database, January 2000

A;Reference number: Z23034  
A;Accession: T46469  
A;Status: preliminary  
A;Molecule type: mRNA  
A;Residues: 1-319 <AAA>  
A;Cross-references: EMBL:AL137483  
A;Experimental source: adult testis; clone DKFZp434G1930  
C;Genetics:  
A;Note: DKFZp434G1930.1

Query Match 18.0%; Score 143.5; DB 2; Length 319;  
Best Local Similarity 32.8%; Pred. No. 1.6e-06;  
Matches 38; Conservative 15; Mismatches 44; Indels 19; Gaps 7;

```
QY      31 YKCGLCAAAQWDFGNTMCQHQRVHGHHSYKCGLCTRSQKEGLHYTCSSHF-----P 83
      ||| | : :: |||:| ||| | :: | |
Db       9 YKCNECEKT-FSHRSSLLSHQRIH-TGEKPYKCNECEKAFSNS--STLIKHLRVHTGEKP 64

QY      84 YSQYQFWKNF---QTLKIHQRVHGGGGSYKCGLCQEFFGLNNCSSSNRLDGHQRVH 136
      | : | | ||:| | | ||| |:: | | : : | ||:|
Db      65 YRCRECGKAFSQCSTLTVHQRIHTGEKLYKCGECEKAF---NCRA--KLHRHQRIH 115
```

RESULT 5

I53869

zinc finger protein - mouse

C;Species: Mus musculus (house mouse)

C;Date: 29-May-1998 #sequence\_revision 29-May-1998 #text\_change 01-Dec-2000

C;Accession: I53869

R;Brady, J.P.; Piatigorsky, J.

Gene 149, 299-304, 1994

A;Title: A mouse cDNA encoding a protein with zinc-fingers and a KRAB domain shows similarity to human profilaggrin.

A;Reference number: I53869; MUID:95047492; PMID:7959006

A;Accession: I53869

A;Status: preliminary; translated from GB/EMBL/DDBJ

A;Molecule type: mRNA

A;Residues: 1-555 <RES>

A;Cross-references: GB:L28167; NID:g758660; PIDN:AAA67545.1; PID:g758661

C;Superfamily: zinc finger protein ZFP-36; LIM metal-binding repeat homology

Query Match 17.7%; Score 141; DB 2; Length 555;  
Best Local Similarity 23.8%; Pred. No. 4.7e-06;  
Matches 50; Conservative 25; Mismatches 47; Indels 88; Gaps 11;

```
QY      4 QVSSPIYDINYYTSEPCQ-----KINVQIAAYKCGLCAAAQWDFGNTMCQH 50
      :| :| : || | | : ||: :: | | | :| : : |
Db     303 RVCTP---VKPYTCEQCEKSLLDVQHLMShVKVHTRE-RPYNCETCGSA-FSQASHLQDH 357

QY     51 QRVH-----GHHHHS-----YKCGLCTRSQKEGLHYTCSSHF- 82
      ||:| | | | ||| | :| : |||:
Db    358 QRLHTGEKPFKCDACGKSFSRSSHLRSHQRVHTGEKPYKCGECGKS-----FICSSNLY 411

QY     83 -----PYSQYQFWKNF---QTLKIHQRVHGGGGSYKCGLCQEFFGLNN----- 122
      || | | :| :| | | | :| : : ||:
Db    412 IHQRVHTGEKPYKCVDCGKEFSRPSLQAHQGIHTGEKSYVCTMCGKGYTLNSNLQVHLR 471
```

Qy 123 -----C-----SSSNRLDGHQRVH 136  
 | | |::| | | |  
 Db 472 VHTGEKPYSCDVCCKGFSRSSQLQSHQRVH 501

RESULT 6

F31201

GLI-related finger protein HKR4 - human (fragments)

C;Species: Homo sapiens (man)

C;Date: 31-Mar-1990 #sequence\_revision 31-Mar-1990 #text\_change 17-Mar-1994

C;Accession: F31201

R;Ruppert, J.M.; Kinzler, K.W.; Wong, A.J.; Bigner, S.H.; Kao, F.T.; Law, M.L.;  
 Seuanez, H.N.; O'Brien, S.J.; Vogelstein, B.

Mol. Cell. Biol. 8, 3104-3113, 1988

A;Title: The GLI-Kruppel family of human genes.

A;Reference number: A93103; MUID:89096896; PMID:2850480

A;Accession: F31201

A;Molecule type: DNA

A;Residues: 1-223 <RUP>

C;Keywords: DNA binding; zinc finger

Query Match 17.4%; Score 138.5; DB 2; Length 223;  
 Best Local Similarity 30.7%; Pred. No. 3.4e-06;  
 Matches 42; Conservative 13; Mismatches 39; Indels 43; Gaps 7;

Qy 31 YKCGLCAAAQWDFGNTMCQHQRVHGHHSYKCGLCTRSQKEGLHYTCSSHFYPYSQYQFW 90  
 |::| | | : : || |::| : |||| | | | : || : ||: |  
 Db 96 YECGQCGRA-FSHSSHFTQHLRIH-NGEKPYKCGECVRHQR--LH---TGEKPYACSQCG 148  
 Qy 91 KNF---QTLKIHQRVHGGGGSYK-----CGLCQEFFG 119  
 | | | | | | | | | | : | | | | : || | : ||  
 Db 149 KAFIWSSVLIHQRIHTGEKPYECSDCGKAFRGRSHFFRHLRTHHTGEKPFACGACGKAFG 208  
 Qy 120 LNNCSSSNRLDGHQRVH 136  
 |::| | | | |  
 Db 209 -----QSSQLIQHQRVH 220

RESULT 7

I38599

zinc finger protein ZNF134 - human

C;Species: Homo sapiens (man)

C;Date: 23-Feb-1996 #sequence\_revision 23-Feb-1996 #text\_change 05-Nov-1999

C;Accession: I38599

R;Tommerup, N.; Vissing, H.

Genomics 27, 259-264, 1995

A;Title: Isolation and fine mapping of 16 novel human zinc finger-encoding cDNAs  
 identify putative candidate genes for developmental and malignant disorders.

A;Reference number: A57785; MUID:96044430; PMID:7557990

A;Accession: I38599

A;Status: preliminary

A;Molecule type: mRNA

A;Residues: 1-348 <RES>

A;Cross-references: EMBL:U09412; NID:g488552; PIDN:AAC50253.1; PID:g488553

C;Genetics:

A;Gene: GDB:ZNF134

A;Cross-references: GDB:137033



A;Map position: 19q13.4-19q13.4

Query Match 17.1%; Score 136.5; DB 2; Length 348;  
Best Local Similarity 27.3%; Pred. No. 8.2e-06;  
Matches 45; Conservative 13; Mismatches 46; Indels 61; Gaps 9;

```
Qy      31 YKCGLCAAAQWDFGNTMCQHQRVHGHHHSYKCGLC-----TRSQKEGLH-----YT 77
      ||| | | : :|: |||:| | | | | | | : :| |
Db      97 YKCSECGKA-FSRKDTLVQHQRHS-GEKPYECSECGKAFSRKATLVQHQRHTGERPYE 154

Qy      78 CS-----SHFPYSQYQFWKNF---QTLKIHQRVHGGSYKCGLC 114
      || | | | | | | | | | | | | | | | | | |
Db     155 CSECGKTFSRKDNLTQHKRIHTGEMPYKCNECGKYFSHSNLIVHQRVHNGARPYKCSDC 214

Qy     115 QEFF-----GLN--NCSSSNRLDG-----HQRVH 136
      : | | | | | | | | | | | | | | | |
Db     215 GKVFRHKSTLVQHESIHTGENPYDCSDCGKSFGHKYTLIKHQRH 259
```

RESULT 8

I39311

Kruppel-type zinc finger protein ZNF74 - human

C;Species: Homo sapiens (man)

C;Date: 06-Sep-1996 #sequence\_revision 06-Sep-1996 #text\_change 01-Dec-2000

C;Accession: I39311; F42825; S37482

R;Aubry, M.; Demczuk, S.; Desmaze, C.; Aikem, M.; Aurias, A.; Julien, J.P.;  
Rouleau, G.A.

Hum. Mol. Genet. 2, 1583-1587, 1993

A;Title: Isolation of a zinc finger gene consistently deleted in DiGeorge  
syndrome.

A;Reference number: I39311; MUID:94093543; PMID:8268910

A;Accession: I39311

A;Status: preliminary; translated from GB/EMBL/DDBJ

A;Molecule type: mRNA

A;Residues: 1-572 <RES>

A;Cross-references: EMBL:X71623; NID:g1050529; PIDN:CAA50632.1; PID:g1050530

A;Note: submitted to the EMBL Data Library, April 1993, revised 01-NOV-1995

R;Aubry, M.; Marineau, C.; Zhang, F.R.; Zahed, L.; Figlewicz, D.; Delattre, O.;  
Thomas, G.; de Jong, P.J.; Julien, J.P.; Rouleau, G.A.  
Genomics 13, 641-648, 1992

A;Title: Cloning of six new genes with zinc finger motifs mapping to short and  
long arms of human acrocentric chromosome 22 (p and q11.2).

A;Reference number: A42825; MUID:92347859; PMID:1639391

A;Accession: F42825

A;Status: preliminary; not compared with conceptual translation

A;Molecule type: nucleic acid

A;Residues: 427-444, 'V', 446-511 <AUB>

A;Note: sequence extracted from NCBI backbone (NCBIP:109775)

C;Superfamily: zinc finger protein ZFP-36; LIM metal-binding repeat homology

Query Match 17.0%; Score 135.5; DB 2; Length 572;  
Best Local Similarity 29.2%; Pred. No. 1.7e-05;  
Matches 33; Conservative 14; Mismatches 25; Indels 41; Gaps 5;

```
Qy      31 YKCGLCAAAQWDFGNTMCQHQRVHGHHH-----SYKCGLCTRSQKEGLHYTCSSHFP 83
      |:|| | | :|| | | | | | | | | | | | |
Db     288 YRCGECGKA-----FNQRTHLTRHHRIHTGEKPYQCGSC-----GKAFTCHS--- 329
```

Qy 84 YSQYQFWKNFQTLKIHQRVHGGGGSYKCGLCQEFFGLNNCSSSNRLDGHQRVH 136  
 :| :|::| | :|| |:: | :| :|| ||| |  
 Db 330 -----SLTVHEKIHSGDKPFKCSDEKAF-----NSRSRLTLHQRT 366

# RESULT 9

I48668

zinc finger protein 51 - mouse

N;Alternate names: finger protein zfec12

C;Species: Mus musculus (house mouse)

C;Date: 02-Jul-1996 #sequence\_revision 02-Jul-1996 #text\_change 08-Dec-2000

C;Accession: I48668; A40984

R;Burke, P.S.; Don, J.; Wolgemuth, D.J.

Mamm. Genome 5, 387-389, 1994

A;Title: Zfp-51, a murine zinc finger encoding gene mapping to the t-complex region of chromosome 17, encodes 19 contiguous zinc fingers and is ubiquitously expressed.

A;Reference number: I48668; MUID:94319090; PMID:8043957

A;Accession: I48668

A;Status: preliminary; translated from GB/EMBL/DDBJ

A;Molecule type: mRNA

A;Residues: 1-710 <RES>

A;Cross-references: EMBL:X74855; NID:g488832; PIDN:CAA52847.1; PID:g488833

R;Crossley, P.H.; Little, P.F.R.

Proc. Natl. Acad. Sci. U.S.A. 88, 7923-7927, 1991

A;Title: A cluster of related zinc finger protein genes is deleted in the mouse embryonic lethal mutation t(w18).

A;Reference number: A40984; MUID:91376058; PMID:1680234

A;Accession: A40984

A;Status: preliminary

A;Molecule type: mRNA

A;Residues: 672-710 <CRO>

A;Cross-references: GB:M74235

C;Genetics:

A;Gene: Zfp-51

C;Superfamily: zinc finger protein ZFP-36; LIM metal-binding repeat homology

C;Keywords: DNA binding; zinc finger

Query Match 16.8%; Score 134; DB 2; Length 710;  
 Best Local Similarity 28.6%; Pred. No. 2.9e-05;  
 Matches 38; Conservative 20; Mismatches 39; Indels 36; Gaps 7;

Qy 9 IYDINYYTS-----EPCQKINVKQIAAYKCGLCAAAQWDFGNTMCQHQRVHGHHSYKC 63  
 :|| :|::| :| | | ||| | :: ::| |::| |||  
 Db 120 VYD-DYFSSAYSLMQPTAYIKEKPHQCCKGKGC---FNTSSSLTVHHRH-TKEKPYKC 173  
 Qy 64 GLCTRSQKEGLHYTCSSHPYSQYQFWKNFQTLKIHQRVHGGGGSYKCGLCQEFFGLNNC 123  
 :| :| :| :| ||||| | | ||| |::|  
 Db 174 SVCDKS-----FTQCTH-----LKIHQRRHTGEKPYKCKEKEKSF----- 208  
 Qy 124 SSSNRLDGHQRVH 136  
 : | ||::|  
 Db 209 VQLSALKSHQKLH 221

# RESULT 10

S03679  
 finger protein (clone mkr5) - mouse (fragment)  
 C;Species: Mus musculus (house mouse)  
 C;Date: 04-Dec-1992 #sequence\_revision 04-Dec-1992 #text\_change 01-Dec-2000  
 C;Accession: S03679  
 R;Chowdhury, K.; Rohdewohld, H.; Gruss, P.  
 Nucleic Acids Res. 16, 9995-10011, 1988  
 A;Title: Specific and ubiquitous expression of different Zn finger protein genes in the mouse.  
 A;Reference number: S03677; MUID:89057528; PMID:3143103  
 A;Accession: S03679  
 A;Status: not compared with conceptual translation  
 A;Molecule type: mRNA  
 A;Residues: 1-475 <CHO>  
 A;Note: the sequence in fig.2 is inconsistent with that shown in fig.1 in having 64-Asn, 159-Trp, 187-trp, 188-trp, and in lacking a Met after 64-Gly and an Asp after 78-Ser  
 C;Superfamily: zinc finger protein ZFP-36; LIM metal-binding repeat homology

Query Match 16.8%; Score 133.5; DB 2; Length 475;  
 Best Local Similarity 27.8%; Pred. No. 2.2e-05;  
 Matches 40; Conservative 24; Mismatches 49; Indels 31; Gaps 8;

```

Qy      15 YTSEPCQK-----INVKQ-----IAAYKCGLCAAQWDFGNTMCQHQRVHGHHSYK 62
      || || | : || | :| :| | : :| ||||| :|
Db      223 YTCEVCHKSFYRGSSLTVHQRIHTGEKPYECEICRKA-FSHHASLTQHQRVHS-GEKPFK 280

Qy      63 CGLCTRSQKEGLHYTCSSHF-----PYSQYQFWKNF---QTLKIHQRVHGGGGSYKCG 112
      | | :: :: :| :||: | : | :| | |||:| | ::|
Db      281 CKECGKAFRQNIH--LASHWRIHTGEKPFECGEGCKSFSISSQLATHQRIHTGEKPFECK 338

Qy      113 LCQEFFGLNNCSSSNRLDGHQRVH 136
      :|:: | | | | | :|
Db      339 VCRKAFRQN-----IHLASHWRIH 357

```

# RESULT 11

I48689  
 gene NK10 protein - mouse  
 C;Species: Mus musculus (house mouse)  
 C;Date: 02-Jul-1996 #sequence\_revision 02-Jul-1996 #text\_change 01-Dec-2000  
 C;Accession: I48689; S49078  
 R;Lange, R.; Christoph, A.; Thiesen, H.J.; Vopper, G.; Johnson, K.R.; Lemaire, L.; Plomann, M.; Cremer, H.; Barthels, D.; Heinleim, U.A.O.  
 DNA Cell Biol. 14, 971-981, 1995  
 A;Title: Developmentally regulated mouse gene NK10 eucodes a xZinc Finger Repressor Protein with differential NDA-Binding Domains.  
 A;Reference number: I48689; MUID:96069544; PMID:7576184  
 A;Accession: I48689  
 A;Status: preliminary; translated from GB/EMBL/DDBJ  
 A;Molecule type: mRNA  
 A;Residues: 1-636 <RES>  
 A;Cross-references: EMBL:X79828; NID:g506501; PIDN:CAA56225.1; PID:g506502  
 C;Superfamily: zinc finger protein ZFP-36; LIM metal-binding repeat homology

Query Match 16.8%; Score 133.5; DB 2; Length 636;  
 Best Local Similarity 31.6%; Pred. No. 2.9e-05;

Matches 37; Conservative 15; Mismatches 44; Indels 21; Gaps 6;

```

Qy      31 YKCGLCAAAQWDFGNTMCQHQRVHGHHHSYKCGLCTRS-----QKEGLHYTCSSHF 82
      |:| || | :  :: |||:|      |:| || ||      | | | |
Db      306 YQCSLCGKA-FQRSSSLVQHQRH-TGEKPYRCNLGSRFRHSTSLTQHEVTH---SGEK 360

Qy      83 PYSQYQFWKNF---QTLKIHQRVHGGGGSYKCGLCQEFFGLNNCSSSNRLDGHQRVH 136
      |:  : | |  :| |:| | |  ::| :|  ||      | | | |:|
Db      361 PFQCKEKGKAFSRCSSLVQHERHTTGEKPFECISICGRAFG-----QSPSLYKHMRIH 412

```

# RESULT 12

T14757

hypothetical protein DKFZp572C163.1 - human (fragment)

C;Species: Homo sapiens (man)

C;Date: 20-Sep-1999 #sequence\_revision 20-Sep-1999 #text\_change 01-Dec-2000

C;Accession: T14757

R;Wambutt, R.; Heubner, D.; Mewes, H.W.; Gassenhuber, J.; Wiemann, S.

submitted to the Protein Sequence Database, August 1999

A;Reference number: Z18181

A;Accession: T14757

A;Status: preliminary

A;Molecule type: mRNA

A;Residues: 1-701 <WAM>

A;Cross-references: EMBL:AL110217

A;Experimental source: adult subthalamic nucleus; clone DKFZp572C163

C;Genetics:

A;Note: DKFZp572C163.1

C;Superfamily: zinc finger protein ZFP-36; LIM metal-binding repeat homology

Query Match 16.8%; Score 133.5; DB 2; Length 701;  
 Best Local Similarity 23.4%; Pred. No. 3.2e-05;  
 Matches 41; Conservative 16; Mismatches 37; Indels 81; Gaps 8;

```

Qy      31 YKCGLCAAAQWDFGNT-----MCQHQRVH-----GHHH-HS- 60
      |:| |      | |      :| |||:|      || |:
Db      468 YECNQC-----GKTFSQORTHLCAHQRIHTGEKPYECNECGKTFADNSALRAHHRIHTG 520

Qy      61 ---YKCGLCTRSQKEGLHY-----TCSSHFYPYSQYQFWKNF----- 93
      |:| | : : | | | | : | |
Db      521 EKPYECNDCGKTFSKTSHLRAHLRTRSGEKPYECSECGKTFSEKSYVSAHQRVHTGEKPY 580

Qy      94 -----QTLKIHQRVHGGGGSYKCGLCQEFFGLNNCSSSNRLDGHQRVH 136
      ||::|||:| | ||:| | : | | : | |||:|
Db      581 ECNVCGKPFPAHNSTLRVHQRIHTGEKSYECNDCGKTF-----SQKSHLSAQRIH 630

```

# RESULT 13

A34612

zinc finger protein ZNF7 - human

N;Alternate names: zinc finger protein kox4

C;Species: Homo sapiens (man)

C;Date: 22-Jun-1990 #sequence\_revision 22-Jun-1990 #text\_change 01-Dec-2000

C;Accession: A34612; A56409; S10421; I37972

R;Lania, L.; Donti, E.; Pannuti, A.; Pascucci, A.; Pengue, G.; Feliciello, I.;

La Mantia, G.; Lanfranccone, L.; Pelicci, P.G.

Genomics 6, 333-340, 1990

A;Title: cDNA isolation, expression analysis, and chromosomal localization of two human zinc finger genes.

A;Reference number: A34612; MUID:90169993; PMID:2106481

A;Accession: A34612

A;Molecule type: mRNA

A;Residues: 1-686 <LAN>

A;Cross-references: GB:M29580; NID:g340445; PIDN:AAA61313.1; PID:g340446; GB:J04751

R;Bray, P.; Lichter, P.; Thiesen, H.J.; Ward, D.C.; Dawid, I.B.

Proc. Natl. Acad. Sci. U.S.A. 88, 9563-9567, 1991

A;Title: Characterization and mapping of human genes encoding zinc finger proteins.

A;Reference number: A56409; MUID:92052132; PMID:1946370

A;Accession: A56409

A;Molecule type: DNA

A;Residues: 425-589 <BRA>

A;Cross-references: GB:M77170

R;Thiesen, H.J.

submitted to the EMBL Data Library, March 1990

A;Reference number: S10397

A;Accession: S10421

A;Molecule type: mRNA

A;Residues: 413-468 <THI>

A;Cross-references: EMBL:X52335; NID:g34165; PIDN:CAA36561.1; PID:g930095

R;Thiesen, H.J.

New Biol. 2, 363-374, 1990

A;Title: Multiple genes encoding zinc finger domains are expressed in human T cells.

A;Reference number: I37949; MUID:91145339; PMID:2288909

A;Accession: I37972

A;Status: translated from GB/EMBL/DDBJ

A;Molecule type: mRNA

A;Residues: 413-468 <RES>

A;Cross-references: EMBL:X52335; NID:g34165; PIDN:CAA36561.1; PID:g930095

C;Genetics:

A;Gene: GDB:ZNF7

A;Cross-references: GDB:120509; OMIM:194531

A;Map position: 8q24.3-8q24.3

C;Superfamily: zinc finger protein ZFP-36; LIM metal-binding repeat homology

C;Keywords: DNA binding; zinc finger

Query Match 16.5%; Score 131.5; DB 2; Length 686;  
Best Local Similarity 26.9%; Pred. No. 4.8e-05;  
Matches 45; Conservative 11; Mismatches 46; Indels 65; Gaps 8;

```
Qy      31 YKCGLCAAAQWDFG--NTMCQHQRVHGHHHSYKCGLCTRSQKEGLH----- 75
      ||| | | || : : :||| | :|| | : :| |
Db      441 YKCNKCTKA---FGCSSRLIRHQORTH-TGEKPFKCECGKGFVQGSHLIQHQRIHTGEKP 496

Qy      76 YTCS-----SHFPYSQYQFWKNFQ---TLKIHQRVHGSGSYKCG 112
      | | : | | | | | | | | | | | | | | | | | | | | | | | | | |
Db      497 YVCNDCGKAFSQSSSLIYHQRIHKGEKPYECLQCGKAFSMSTQLTIHQRVHTGERPYKCN 556

Qy     113 LCQEFFGLNN-----CSS-----SNRLDGHQVRVH 136
      | : | | : | | | | | | | | | | | | | | | | | | | | | |
Db     557 ECGKAFSQNSTLFQHQIIHAGVKPYECSECGKAFSRSSYLIEHQRIH 603
```

# RESULT 14

S70006

finger protein zfOC1 - guinea pig

C;Species: Cavia porcellus (guinea pig)

C;Date: 06-Dec-1996 #sequence\_revision 13-Mar-1997 #text\_change 05-Nov-1999

C;Accession: S70006

R;Rivolta, M.N.; Negrini, C.; Wilcox, E.R.

Biochim. Biophys. Acta 1306, 127-132, 1996

A;Title: A novel zinc finger gene preferentially expressed in the retina and the organ of Corti localizes to human chromosome 12q24.3.

A;Reference number: S70006; MUID:96221281; PMID:8634327

A;Accession: S70006

A;Status: preliminary

A;Molecule type: mRNA

A;Residues: 1-261 <RIV>

A;Cross-references: EMBL:L26335; NID:g1237277; PIDN:AAC42091.1; PID:g1237278

C;Keywords: zinc

Query Match 16.4%; Score 131; DB 2; Length 261;

Best Local Similarity 28.8%; Pred. No. 2.1e-05;

Matches 42; Conservative 16; Mismatches 44; Indels 44; Gaps 7;

```

Qy      26 KQIAAYKCG-----LCAAAQWDFGNT--MCQHQRVHGHHSY 61
      |   |::|               ::   | | : | |||||   :
Db      85 KPYKCYECGKAFNWSPHLQIHMRVHTGEKPYVCSECGRGFSNSSLNLCMHQRVH-TGEKPF 143

Qy      62 KCGLCTRS-----QKEGLHYTCSSHPYSQYQFWKNF---QTLKIHQRVHGGGGSYK 110
      || | ::   : :|   :   || | : | | : | ||||| | | :
Db     144 KCEECGKAFRHTSSSLCMHQRVH---TGEKPYKCYECGKAFSQSSSLCIHQRVHTGEKPYR 200

Qy     111 CGLCQEFFGLNNCSSSNRLDGHQRVH 136
      | | : |   | | : |   |||||
Db     201 CCGCGKAF-----SQSSSLCIHQRVH 221

```

# RESULT 15

JH0500

zinc finger protein 29 - mouse

C;Species: Mus musculus (house mouse)

C;Date: 31-Mar-1992 #sequence\_revision 31-Mar-1992 #text\_change 08-Dec-2000

C;Accession: JH0500

R;Denny, P.; Ashworth, A.

Gene 106, 221-227, 1991

A;Title: A zinc finger protein-encoding gene expressed in the post-meiotic phase of spermatogenesis.

A;Reference number: JH0500; MUID:92039080; PMID:1937051

A;Accession: JH0500

A;Molecule type: mRNA

A;Residues: 1-614 <DEN>

A;Cross-references: GB:X55126; NID:g55470; PIDN:CAA38920.1; PID:g55471

A;Experimental source: testis

C;Keywords: DNA binding; zinc finger

F;217-244/Region: zinc finger

F;245-272/Region: zinc finger

F;273-300/Region: zinc finger

F;301-328/Region: zinc finger

F;329-356/Region: zinc finger  
 F;357-384/Region: zinc finger  
 F;385-412/Region: zinc finger  
 F;413-440/Region: zinc finger  
 F;441-468/Region: zinc finger  
 F;469-496/Region: zinc finger  
 F;497-524/Region: zinc finger  
 F;525-552/Region: zinc finger  
 F;553-580/Region: zinc finger  
 F;581-608/Region: zinc finger

Query Match 16.2%; Score 129.5; DB 2; Length 614;  
 Best Local Similarity 30.3%; Pred. No. 6.8e-05;  
 Matches 37; Conservative 12; Mismatches 42; Indels 31; Gaps 6;

```

Qy      31 YKCGLCAAAQWDFGNTMCQHQRVHGHHSYKCGLCTRS-----QKEGLH----- 75
      |:| | : : : |:| | ||||| :| :| |
Db      418 YQCGEC-GKNFSRSSNLATHRRTH-LVEKPYKCGLCGKSFSQSSSLIAHQGTHTGEKPYE 475

Qy      76 -YTCSSHFPYSQYQFWKNFQTLKIHQRVHGGGGSYKCGLCQEFFGLNNCSSNRLDGHQR 134
      || | :| | ||| | | |:| | : | | :| |||
Db      476 CLTCGESFSWS-----SNLIKHQRTHTGEKPYRCGDCGKGF-----SQRSQLVVHQR 522

Qy      135 VH 136
      |
Db      523 TH 524
  
```

Search completed: September 17, 2004, 16:49:56  
 Job time : 23.4542 secs

OM protein - protein search, using sw model

Run on: September 17, 2004, 16:48:23 ; Search time 83.7356 Seconds  
(without alignments)  
529.242 Million cell updates/sec

Title: US-10-057-890A-10  
Perfect score: 797  
Sequence: 1 MDYQVSSPIYDINYYTSEPC.....GLNNCSSSNRLDGHQRVHAA 138

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 1342398 seqs, 321133274 residues

Total number of hits satisfying chosen parameters: 1342398

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Published\_Applications\_AA:\*

- 1: /cgn2\_6/ptodata/2/pubpaa/US07\_PUBCOMB.pep:\*
- 2: /cgn2\_6/ptodata/2/pubpaa/PCT\_NEW\_PUB.pep:\*
- 3: /cgn2\_6/ptodata/2/pubpaa/US06\_NEW\_PUB.pep:\*
- 4: /cgn2\_6/ptodata/2/pubpaa/US06\_PUBCOMB.pep:\*
- 5: /cgn2\_6/ptodata/2/pubpaa/US07\_NEW\_PUB.pep:\*
- 6: /cgn2\_6/ptodata/2/pubpaa/PCTUS\_PUBCOMB.pep:\*
- 7: /cgn2\_6/ptodata/2/pubpaa/US08\_NEW\_PUB.pep:\*
- 8: /cgn2\_6/ptodata/2/pubpaa/US08\_PUBCOMB.pep:\*
- 9: /cgn2\_6/ptodata/2/pubpaa/US09A\_PUBCOMB.pep:\*
- 10: /cgn2\_6/ptodata/2/pubpaa/US09B\_PUBCOMB.pep:\*
- 11: /cgn2\_6/ptodata/2/pubpaa/US09C\_PUBCOMB.pep:\*
- 12: /cgn2\_6/ptodata/2/pubpaa/US09\_NEW\_PUB.pep:\*
- 13: /cgn2\_6/ptodata/2/pubpaa/US10A\_PUBCOMB.pep:\*
- 14: /cgn2\_6/ptodata/2/pubpaa/US10B\_PUBCOMB.pep:\*
- 15: /cgn2\_6/ptodata/2/pubpaa/US10C\_PUBCOMB.pep:\*
- 16: /cgn2\_6/ptodata/2/pubpaa/US10\_NEW\_PUB.pep:\*
- 17: /cgn2\_6/ptodata/2/pubpaa/US60\_NEW\_PUB.pep:\*
- 18: /cgn2\_6/ptodata/2/pubpaa/US60\_PUBCOMB.pep:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

8

Result	Query					
No.	Score	Match	Length	DB	ID	Description



1	797	100.0	138	14	US-10-057-890A-10	Sequence 10, Appl
2	797	100.0	157	14	US-10-057-890A-31	Sequence 31, Appl
3	379	47.6	352	9	US-09-725-285-2	Sequence 2, Appli
4	379	47.6	352	9	US-09-759-841-2	Sequence 2, Appli
5	379	47.6	352	9	US-09-779-879A-22	Sequence 22, Appl
6	379	47.6	352	9	US-09-779-880A-22	Sequence 22, Appl
7	379	47.6	352	9	US-09-813-653-15	Sequence 15, Appl
8	379	47.6	352	9	US-09-813-653-17	Sequence 17, Appl
9	379	47.6	352	9	US-09-796-202-1	Sequence 1, Appli
10	379	47.6	352	9	US-09-195-662A-2	Sequence 2, Appli
11	379	47.6	352	9	US-09-339-912A-2	Sequence 2, Appli
12	379	47.6	352	9	US-09-938-719-5	Sequence 5, Appli
13	379	47.6	352	9	US-09-939-226-5	Sequence 5, Appli
14	379	47.6	352	9	US-09-938-703-5	Sequence 5, Appli
15	379	47.6	352	9	US-09-502-783A-2	Sequence 2, Appli
16	379	47.6	352	10	US-09-734-221A-14	Sequence 14, Appl
17	379	47.6	352	11	US-09-826-509-477	Sequence 477, App
18	379	47.6	352	13	US-10-106-623-2	Sequence 2, Appli
19	379	47.6	352	14	US-10-232-686-2	Sequence 2, Appli
20	379	47.6	352	14	US-10-086-814-1	Sequence 1, Appli
21	379	47.6	352	14	US-10-067-800-22	Sequence 22, Appl
22	379	47.6	352	14	US-10-290-058A-6	Sequence 6, Appli
23	379	47.6	352	14	US-10-225-567A-352	Sequence 352, App
24	379	47.6	352	14	US-10-323-314-1	Sequence 1, Appli
25	379	47.6	352	14	US-10-072-301-1	Sequence 1, Appli
26	379	47.6	352	14	US-10-071-866-1	Sequence 1, Appli
27	379	47.6	352	14	US-10-135-839-22	Sequence 22, Appl
28	379	47.6	352	14	US-10-239-423-67	Sequence 67, Appl
29	379	47.6	352	14	US-10-439-845-4	Sequence 4, Appli
30	379	47.6	352	15	US-10-360-828-1	Sequence 1, Appli
31	379	47.6	352	16	US-10-661-798-5	Sequence 5, Appli
32	379	47.6	352	16	US-10-791-905-2	Sequence 2, Appli
33	379	47.6	352	16	US-10-612-791-5	Sequence 5, Appli
34	374	46.9	352	14	US-10-164-649-52	Sequence 52, Appl
35	374	46.9	352	14	US-10-439-845-2	Sequence 2, Appli
36	373	46.8	352	9	US-09-779-879A-2	Sequence 2, Appli
37	373	46.8	352	9	US-09-779-880A-2	Sequence 2, Appli
38	373	46.8	352	14	US-10-067-800-2	Sequence 2, Appli
39	373	46.8	352	14	US-10-135-839-2	Sequence 2, Appli
40	363	45.5	352	13	US-10-106-623-20	Sequence 20, Appl
41	359	45.0	352	12	US-10-151-274-5	Sequence 5, Appli
42	258	32.4	184	9	US-09-938-719-4	Sequence 4, Appli
43	258	32.4	184	9	US-09-939-226-4	Sequence 4, Appli
44	258	32.4	184	9	US-09-938-703-4	Sequence 4, Appli
45	258	32.4	184	16	US-10-661-798-4	Sequence 4, Appli

# ALIGNMENTS

RESULT 1  
US-10-057-890A-10  
; Sequence 10, Application US/10057890A  
; Publication No. US20030044901A1  
; GENERAL INFORMATION:  
; APPLICANT: Coleman, Timothy

```
; APPLICANT: Mansfield, Brian
; TITLE OF INVENTION: Scaffold Fusion Polypeptides, Composition for Making the
Same, and Methods
; TITLE OF INVENTION: of Using the Same.
; FILE REFERENCE: PF537
; CURRENT APPLICATION NUMBER: US/10/057,890A
; CURRENT FILING DATE: 2002-01-29
; PRIOR APPLICATION NUMBER: 60/265,782
; PRIOR FILING DATE: 2001-01-31
; PRIOR APPLICATION NUMBER: 60/265,858
; PRIOR FILING DATE: 2001-01-31
; NUMBER OF SEQ ID NOS: 32
; SEQ ID NO 10
; LENGTH: 138
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-057-890A-10
```

```
Query Match          100.0%; Score 797; DB 14; Length 138;
Best Local Similarity 100.0%; Pred. No. 5.6e-71;
Matches 138; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
Qy      1 MDYQVSSPIYDINYYTSEPCQKINVKQIAAYKCGLCAAQWDFGNTMCQHQRVHGHHS 60
          |||
Db      1 MDYQVSSPIYDINYYTSEPCQKINVKQIAAYKCGLCAAQWDFGNTMCQHQRVHGHHS 60

Qy      61 YKGLCTRSQKEGLHYTCSSHFYPYSQYQFWKNFQTLKIHQRVHGGGGSYKGLCQEFFGL 120
          |||
Db      61 YKGLCTRSQKEGLHYTCSSHFYPYSQYQFWKNFQTLKIHQRVHGGGGSYKGLCQEFFGL 120

Qy      121 NNCSSSNRLDGHQRVHAA 138
          |||
Db      121 NNCSSSNRLDGHQRVHAA 138
```

# RESULT 2

US-10-057-890A-31

```
; Sequence 31, Application US/10057890A
; Publication No. US20030044901A1
; GENERAL INFORMATION:
; APPLICANT: Coleman, Timothy
; APPLICANT: Mansfield, Brian
; TITLE OF INVENTION: Scaffold Fusion Polypeptides, Composition for Making the
Same, and Methods
; TITLE OF INVENTION: of Using the Same.
; FILE REFERENCE: PF537
; CURRENT APPLICATION NUMBER: US/10/057,890A
; CURRENT FILING DATE: 2002-01-29
; PRIOR APPLICATION NUMBER: 60/265,782
; PRIOR FILING DATE: 2001-01-31
; PRIOR APPLICATION NUMBER: 60/265,858
; PRIOR FILING DATE: 2001-01-31
; NUMBER OF SEQ ID NOS: 32
; SEQ ID NO 31
; LENGTH: 157
; TYPE: PRT
; ORGANISM: Homo sapiens
```

US-10-057-890A-31

Query Match 100.0%; Score 797; DB 14; Length 157;  
Best Local Similarity 100.0%; Pred. No. 6.4e-71;  
Matches 138; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy      1 MDYQVSSPIYDINYYTSEPCQKINVKQIAAYKCGLCAAQWDFGNTMCQHQRVHGHHS 60
          |||
Db      20 MDYQVSSPIYDINYYTSEPCQKINVKQIAAYKCGLCAAQWDFGNTMCQHQRVHGHHS 79

Qy      61 YKGLCTRSQKEGLHYTCSSHPYSQYQFWKNFQTLKIHQRVHGGGGSYKCGLCQEFFGL 120
          |||
Db      80 YKGLCTRSQKEGLHYTCSSHPYSQYQFWKNFQTLKIHQRVHGGGGSYKCGLCQEFFGL 139

Qy      121 NNCSSSNRLDGHQRVHAA 138
          |||
Db      140 NNCSSSNRLDGHQRVHAA 157
```

RESULT 3

US-09-725-285-2

; Sequence 2, Application US/09725285

; Patent No. US20010000241A1

; GENERAL INFORMATION:

; APPLICANT: Li, Yi

; APPLICANT: Ruben, Steven, M.

; TITLE OF INVENTION: Antibodies to Human G-Protein Chemokine Receptor  
HDGMR10

; TITLE OF INVENTION: (CCR5 Receptor)

; FILE REFERENCE: 1488.1150003

; CURRENT APPLICATION NUMBER: US/09/725,285

; CURRENT FILING DATE: 2000-11-29

; PRIOR APPLICATION NUMBER: 09/339,912

; PRIOR FILING DATE: 1999-06-25

; PRIOR APPLICATION NUMBER: 09/195,662

; PRIOR FILING DATE: 1998-11-18

; PRIOR APPLICATION NUMBER: 08/466,343

; PRIOR FILING DATE: 1995-06-06

; NUMBER OF SEQ ID NOS: 9

; SOFTWARE: PatentIn version 3.0

; SEQ ID NO 2

; LENGTH: 352

; TYPE: PRT

; ORGANISM: Artificial Sequence: Genomic

; FEATURE:

; OTHER INFORMATION: Deduced Amino Acid Sequence

US-09-725-285-2

Query Match 47.6%; Score 379; DB 9; Length 352;  
Best Local Similarity 35.0%; Pred. No. 3e-29;  
Matches 99; Conservative 7; Mismatches 27; Indels 150; Gaps 7;

```
Qy      1 MDYQVSSPIYDINYYTSEPCQKINVKQIAA----- 30
          |||
Db      1 MDYQVSSPIYDINYYTSEPCQKINVKQIAARLLPPLYSLVFIFGFVGNMLVILILINCKR 60

Qy      31 -----YKGLC-----AAAQWDFGNTMCQHQRVHGHHSYKCG--- 64
```

Qy	1	MDYQVSSPIYDINYYTSEPCQKINVKQIAA-----	30
Db	1	MDYQVSSPIYDINYYTSEPCQKINVKQIAARLLPPLYSLVFIFGFVGNMLVILILINCKR	60
Qy	31	-----YKCGLC-----AAAQWDFGNTMCQHQRVHGHHSYKCG---	64
		:   :	
Db	61	LKSMTDIYLLNLAISDLFFLLTVPFWAHYAAAQWDFGNTMC--QLLTGLYFIGFFSGIFF	118
Qy	65	-----LCTRSQKEGLHYTC	78
		:	
Db	119	IILLTIDRYLAVVHAVEFALKARTVTFGVVTSVITWVAVFASLPGLIIFTRSQKEGLHYTC	178

```

Qy      79 SSHFPYSQYQFWKNFQTLKI-----HQRVHGG----- 105
      |||||||||||||||||||
Db      179 SSHFPYSQYQFWKNFQTLKIVILGLVLP LLVMVICYSGILKTLLRCRNEKKRHRVRLIF 238

Qy      106 -----GGSYKCGLC-----QEFFGLNNC SSSNRLDGHQRV 135
      | | ||||||||||||||| :|
Db      239 TIMIVYFLFWAPYNIVLLLNTFQEFFGLNNC SSSNRLDQAMQV 281

```

RESULT 5

```

US-09-779-879A-22
; Sequence 22, Application US/09779879A
; Patent No. US20020048786A1
; GENERAL INFORMATION:
; APPLICANT: Rosen, Craig A.
; APPLICANT: Roschke, Viktor
; APPLICANT: Li, Yi
; APPLICANT: Ruben, Steven, M.
; TITLE OF INVENTION: Human G-protein Chemokine Receptor (CCR5) HDG NR10
; FILE REFERENCE: 1488.115000A
; CURRENT APPLICATION NUMBER: US/09/779,879A
; CURRENT FILING DATE: 2001-02-09
; PRIOR APPLICATION NUMBER: US 60/181,258
; PRIOR FILING DATE: 2000-02-09
; PRIOR APPLICATION NUMBER: US 60/187,999
; PRIOR FILING DATE: 2000-03-09
; PRIOR APPLICATION NUMBER: US 60/234,336
; PRIOR FILING DATE: 2000-09-22
; NUMBER OF SEQ ID NOS: 58
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 22
; LENGTH: 352
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-779-879A-22

```

```

Query Match          47.6%; Score 379; DB 9; Length 352;
Best Local Similarity 35.0%; Pred. No. 3e-29;
Matches 99; Conservative 7; Mismatches 27; Indels 150; Gaps 7;

```

```

Qy      1 MDYQVSSPIYDINYYTSEPCQKINVKQIAA----- 30
      |||||||||||||||||||
Db      1 MDYQVSSPIYDINYYTSEPCQKINVKQIAARLLPPLYSLVFIFGFVGNMLVILILINCKR 60

Qy      31 -----YKCGLC-----AAAQWDFGNTMCQHQRVHGH HHHHSYKCG--- 64
      | | ||||||||||| | : | : |
Db      61 LKSMTDIYLLNLAISDLFFLLTVPFWAHYAAAQWDFGNTMC--QLLTGLYFIGFFSGIFF 118

Qy      65 -----LCTRSQKEGLHYTC 78
      : |||||||||||
Db      119 IILLTIDRYLAVVHAVFALKARTVTFGVVT SVITWVAVFASLPGIIFTRSQKEGLHYTC 178

Qy      79 SSHFPYSQYQFWKNFQTLKI-----HQRVHGG----- 105
      |||||||||||||||||||
Db      179 SSHFPYSQYQFWKNFQTLKIVILGLVLP LLVMVICYSGILKTLLRCRNEKKRHRVRLIF 238

```

Qy 106 -----GGSYKCGLC-----QEFFGLNNCSSSNRLDGHQRV 135  
 | | ||||| :|  
 Db 239 TIMIVYFLFWAPYNIVLLLNTFQEFFGLNNCSSSNRLDQAMQV 281

RESULT 6

US-09-779-880A-22  
 ; Sequence 22, Application US/09779880A  
 ; Patent No. US20020061834A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Rosen, Craig A.  
 ; APPLICANT: Roschke, Viktor  
 ; APPLICANT: Li, Yi  
 ; APPLICANT: Ruben, Steven, M.  
 ; TITLE OF INVENTION: Human G-protein Chemokine Receptor (CCR5) HDGNR10  
 ; FILE REFERENCE: 1488.115000C  
 ; CURRENT APPLICATION NUMBER: US/09/779,880A  
 ; CURRENT FILING DATE: 2001-02-09  
 ; PRIOR APPLICATION NUMBER: US 60/181,258  
 ; PRIOR FILING DATE: 2000-02-09  
 ; PRIOR APPLICATION NUMBER: US 60/187,999  
 ; PRIOR FILING DATE: 2000-03-09  
 ; PRIOR APPLICATION NUMBER: US 60/234,336  
 ; PRIOR FILING DATE: 2000-09-22  
 ; NUMBER OF SEQ ID NOS: 58  
 ; SOFTWARE: PatentIn version 3.0  
 ; SEQ ID NO 22  
 ; LENGTH: 352  
 ; TYPE: PRT  
 ; ORGANISM: Homo sapiens  
 US-09-779-880A-22

Query Match 47.6%; Score 379; DB 9; Length 352;  
 Best Local Similarity 35.0%; Pred. No. 3e-29;  
 Matches 99; Conservative 7; Mismatches 27; Indels 150; Gaps 7;

Qy 1 MDYQVSSPIYDINYYTSEPCQKINVKQIAA----- 30  
 |||||  
 Db 1 MDYQVSSPIYDINYYTSEPCQKINVKQIAARLLPLYSLVFIFGFVGNMLVILILINCKR 60  
 Qy 31 -----YKCGLC-----AAAQWDFGNTMCQHQRVHGHHHSYKCG--- 64  
 | | ||||| :| :| :|  
 Db 61 LKSMTDIYLLNLAISDLFFLLTVPFWAHYAAAQWDFGNTMC--QLLTGLYFIGFFSGIFF 118  
 Qy 65 -----LCTRSQKEGLHYTC 78  
 : |||||  
 Db 119 IILLTIDRYLAVVHAVFALKARTVTFGVVTSVITWVAVFASLPGIIFTRSQKEGLHYTC 178  
 Qy 79 SSHFPYSQYQFWKNFQTLKI-----HQRVHGG----- 105  
 ||||| :|  
 Db 179 SSHFPYSQYQFWKNFQTLKIVILGLVLP LLMVICYSGILKTLRCRNEKKRHRAVRLIF 238  
 Qy 106 -----GGSYKCGLC-----QEFFGLNNCSSSNRLDGHQRV 135  
 | | ||||| :|  
 Db 239 TIMIVYFLFWAPYNIVLLLNTFQEFFGLNNCSSSNRLDQAMQV 281

# RESULT 7

US-09-813-653-15

; Sequence 15, Application US/09813653

; Patent No. US20020064770A1

; GENERAL INFORMATION:

; APPLICANT: Nestor, John

; APPLICANT: Wilson, Carol

; APPLICANT: See, Raymond

; APPLICANT: Tan Hehir, Christina

; TITLE OF INVENTION: Binding Compounds and Methods For Identifying Binding Compounds

; FILE REFERENCE: CNS-005

; CURRENT APPLICATION NUMBER: US/09/813,653

; CURRENT FILING DATE: 2001-03-20

; PRIOR APPLICATION NUMBER: US 60/190,946

; PRIOR FILING DATE: 2000-03-21

; PRIOR APPLICATION NUMBER: US 60/190,996

; PRIOR FILING DATE: 2000-03-21

; PRIOR APPLICATION NUMBER: US 60/191,299

; PRIOR FILING DATE: 2000-03-21

; NUMBER OF SEQ ID NOS: 44

; SOFTWARE: PatentIn version 3.0

; SEQ ID NO 15

; LENGTH: 352

; TYPE: PRT

; ORGANISM: Homo sapiens

US-09-813-653-15

Query Match 47.6%; Score 379; DB 9; Length 352;

Best Local Similarity 35.0%; Pred. No. 3e-29;

Matches 99; Conservative 7; Mismatches 27; Indels 150; Gaps 7;

```

Qy      1 MDYQVSSPIYDINYYTSEPCQKINVKQIAA----- 30
          |||
Db      1 MDYQVSSPIYDINYYTSEPCQKINVKQIAARLLPPLYSLVFIFGFVGNMLVILILINCKR 60

Qy      31 -----YKCGLC-----AAAQWDFGNTMCQHQRVHGHHHSYKCG--- 64
          |  |
Db      61 LKSMTDIYLLNLAISDLFFLLTVPFWAHYAAAQWDFGNTMC--QLLTGLYFIGFFSGIFF 118

Qy      65 -----LCTRSQKEGLHYTC 78
          : |||
Db      119 IILLTIDRYLAVVHAVFALKARTVTFGVVTSVITWVAVFASLPGIIFTRSQKEGLHYTC 178

Qy      79 SSHFPYSQYQFWKNFQTLKI-----HQRVHGG----- 105
          |||
Db      179 SSHFPYSQYQFWKNFQTLKIVILGLVLP LLVMVICYSGILKTLRLCRNEKKRHRVRLIF 238

Qy      106 -----GGSYKCGLC----QEFGGLNNCSSSNRLDGHQRV 135
          |  |
Db      239 TIMIVYFLFWAPYNIVLLLNTFQEFFGLNNCSSSNRLDQAMQV 281

```

# RESULT 8

US-09-813-653-17

; Sequence 17, Application US/09813653

; Patent No. US20020064770A1

```
; GENERAL INFORMATION:
; APPLICANT: Nestor, John
; APPLICANT: Wilson, Carol
; APPLICANT: See, Raymond
; APPLICANT: Tan Hehir, Christina
; TITLE OF INVENTION: Binding Compounds and Methods For Identifying Binding
Compounds
; FILE REFERENCE: CNS-005
; CURRENT APPLICATION NUMBER: US/09/813,653
; CURRENT FILING DATE: 2001-03-20
; PRIOR APPLICATION NUMBER: US 60/190,946
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: US 60/190,996
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: US 60/191,299
; PRIOR FILING DATE: 2000-03-21
; NUMBER OF SEQ ID NOS: 44
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 17
; LENGTH: 352
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-813-653-17
```

```
Query Match          47.6%; Score 379; DB 9; Length 352;
Best Local Similarity 35.0%; Pred. No. 3e-29;
Matches 99; Conservative 7; Mismatches 27; Indels 150; Gaps 7;
```

```
Qy      1 MDYQVSSPIYDINYYTSEPCQKINVKQIAA----- 30
          |||
Db      1 MDYQVSSPIYDINYYTSEPCQKINVKQIAARLLPPLYSLVFIFGFVGNMLVILIQINCKR 60

Qy     31 -----YKCGLC-----AAAQWDFGNTMCQHQRVHGHHHSYKCG--- 64
          | |
Db     61 LKSMTDIYLLNLAISDLFFLLTVPFWAHYAAAQWDFGNTMC--QLLTGLYFIGFFSGIFF 118

Qy     65 -----LCTRSQKEGLHYTC 78
          : |||
Db    119 IILLTIDRYLAVVHAVFALKARTVTFGVVTSVITWVAVFASLPGIIFTRSQKEGLHYTC 178

Qy     79 SSHFPYSQYQFWKNFQTLKI-----HQRVHGG----- 105
          |||
Db    179 SSHFPYSQYQFWKNFQTLKIVILGLVLP LLVMVICYSGILKTLRCRNEKKRHRVRLIF 238

Qy    106 -----GGSYKCGLC----QEFGFLNNCSSSNRLDGHQRV 135
          | |
Db    239 TIMIVYFLEWAPYNIVLLLNTFQEFGFLNNCSSSNRLDQAMQV 281
```

# RESULT 9

```
US-09-796-202-1
; Sequence 1, Application US/09796202
; Patent No. US20020068813A1
; GENERAL INFORMATION:
; APPLICANT: Dragic, Tatjana
; APPLICANT: Olson, William
; TITLE OF INVENTION: SULFATED CCR5 PEPTIDES FOR HIV-1 INFECTION
```



; FILE REFERENCE: 2048/61010/JPW/SHS  
 ; CURRENT APPLICATION NUMBER: US/09/796,202  
 ; CURRENT FILING DATE: 2001-02-28  
 ; NUMBER OF SEQ ID NOS: 17  
 ; SOFTWARE: PatentIn version 3.0  
 ; SEQ ID NO 1  
 ; LENGTH: 352  
 ; TYPE: PRT  
 ; ORGANISM: human  
 US-09-796-202-1

Query Match 47.6%; Score 379; DB 9; Length 352;  
 Best Local Similarity 35.0%; Pred. No. 3e-29;  
 Matches 99; Conservative 7; Mismatches 27; Indels 150; Gaps 7;

```

Qy      1 MDYQVSSPIYDINYYTSEPCQKINVKQIAA----- 30
          |||
Db      1 MDYQVSSPIYDINYYTSEPCQKINVKQIAARLLPPLYSLVFIFGFVGNMLVILILINCKR 60

Qy      31 -----YKCGLC-----AAAQWDFGNTMCQHQRVHGHHHSYKCG--- 64
          | |
Db      61 LKSMTDIYLLNLAISDLFFLLTVPEWAHYAAAQWDFGNTMC--QLLTGLYFIGFFSGIFF 118

Qy      65 -----LCTRSQKEGLHYTC 78
          : |||
Db      119 IILLTIDRYLAVVHAVFALKARTVTFGVVTSVITWVAVFASLPGIIFTRSQKEGLHYTC 178

Qy      79 SSHFPYSQYQFWKNFQTLKI-----HQRVHGG----- 105
          |||
Db      179 SSHFPYSQYQFWKNFQTLKIVILGLVLP LLVMVICYSGLKTLRLCRNEKKRHRVRLIF 238

Qy      106 -----GGSYKCGLC----QEFGFLNNCSSSNRLDGHQRV 135
          | |
Db      239 TIMIVYFLFWAPYNIVLLLNTFQEFFGLNNCSSSNRLDQAMQV 281
  
```

RESULT 10

US-09-195-662A-2

; Sequence 2, Application US/09195662A  
 ; Patent No. US20020076745A1

; GENERAL INFORMATION:

; APPLICANT: Li, Yi

; APPLICANT: Ruben, Steven, M.

; TITLE OF INVENTION: Human G-Protein Chemokine Receptor HDGMR10 (CCR5 Receptor)

; FILE REFERENCE: 1488.1150002

; CURRENT APPLICATION NUMBER: US/09/195,662A

; CURRENT FILING DATE: 1998-11-18

; PRIOR APPLICATION NUMBER: 08/466,343

; PRIOR FILING DATE: 1995-06-06

; NUMBER OF SEQ ID NOS: 9

; SOFTWARE: PatentIn version 3.0

; SEQ ID NO 2

; LENGTH: 352

; TYPE: PRT

; ORGANISM: Artificial Sequence: Genomic

; FEATURE:

; OTHER INFORMATION: Deduced Amino Acid Sequence  
US-09-195-662A-2

```
Query Match          47.6%;  Score 379;  DB 9;  Length 352;
Best Local Similarity 35.0%;  Pred. No. 3e-29;
Matches   99;  Conservative   7;  Mismatches  27;  Indels  150;  Gaps    7;

Qy          1 MDYQVSSPIYDINYYTSEPCQKINVKQIAA----- 30
              |||
Db          1 MDYQVSSPIYDINYYTSEPCQKINVKQIAARLLPPLYSLVFIFGFVGNMLVILILINCKR 60

Qy          31 -----YKCGLC-----AAAQWDFGNTMCQHQRVHGHHSYKCG--- 64
              | |
Db          61 LKSMTDIYLLNLAIISDLFFLLTVPFWAHYAAAQWDFGNTMC--QLLTGLYFIGFFSGIFF 118

Qy          65 -----LCTRSQKEGLHYTC 78
              : |||
Db          119 IILLTIDRYLAVVHAVFALKARTVTFGVVTSVITWVAVFASLPGIIFTRSQKEGLHYTC 178

Qy          79 SSHFPYSQYQFWKNFQTLKI-----HQRVHGG----- 105
              |||
Db          179 SSHFPYSQYQFWKNFQTLKIVILGLVLPPLVMVICYSGILKTLLRCRNEKKRHRVRLIF 238

Qy          106 -----GGSYKCGLC----QEFFGLNNCSSSNRLDGHQRV 135
              | |
Db          239 TIMIVYFLFWAPYNIVLLLNTFQEFFGLNNCSSSNRLDQAMQV 281
```

RESULT 11

US-09-339-912A-2  
; Sequence 2, Application US/09339912A  
; Patent No. US20020099176A1  
; GENERAL INFORMATION:  
; APPLICANT: Li, Yi  
; APPLICANT: Ruben, Steven, M.  
; TITLE OF INVENTION: Antibodies to Human G-Protein Chemokine Receptor  
HDGNR10  
; TITLE OF INVENTION: (CCR5 Receptor)  
; FILE REFERENCE: 1488.1150003  
; CURRENT APPLICATION NUMBER: US/09/339,912A  
; CURRENT FILING DATE: 1999-06-25  
; PRIOR APPLICATION NUMBER: 09/195,662  
; PRIOR FILING DATE: 1998-11-18  
; PRIOR APPLICATION NUMBER: 08/466,343  
; PRIOR FILING DATE: 1995-06-06  
; NUMBER OF SEQ ID NOS: 9  
; SOFTWARE: PatentIn version 3.0  
; SEQ ID NO 2  
; LENGTH: 352  
; TYPE: PRT  
; ORGANISM: Artificial Sequence: Genomic  
; FEATURE:  
; OTHER INFORMATION: Deduced Amino Acid Sequence  
US-09-339-912A-2

Query Match 47.6%; Score 379; DB 9; Length 352;  
Best Local Similarity 35.0%; Pred. No. 3e-29;

Matches 99; Conservative 7; Mismatches 27; Indels 150; Gaps 7;

```

Qy      1 MDYQVSSPIYDINYYTSEPCQKINVKQIAA----- 30
          ||||||||||||||||||||||||||||
Db      1 MDYQVSSPIYDINYYTSEPCQKINVKQIAARLLPPLYSLVFIFGFVGNMLVILILINCKR 60

Qy     31 -----YKCGLC-----AAAQWDFGNTMCQHQRVHGHHHSYKCG--- 64
          |  |                      |||||  | : | : |
Db     61 LKSMTDIYLLNLAISDLFFLLTVPFWAHYAAAQWDFGNTMC--QLLTGLYFIGFFSGIFF 118

Qy     65 -----LCTRSQKEGLHYTC 78
          : |||||||
Db    119 IILLTIDRYLAVVHAVFALKARTVTFGVVTSVITWVAVFASLPGIIFTRSQKEGLHYTC 178

Qy     79 SSHFPYSQYQFWKNFQTLKI-----HQRVHGG----- 105
          |||||||||||||||||  :: |
Db    179 SSHFPYSQYQFWKNFQTLKIVILGLVLP LLVMVICYS GILKTL LRCRNEKKRHRVRLIF 238

Qy    106 -----GGSYKCGLC---QEFGNLCSSSNRLDGHQRV 135
          |  |  |||||||||||||  :|
Db    239 TIMIVYFLFWAPYNIVLLNLTFFQEFFGLNLCSSSNRLDQAMQV 281

```

RESULT 12

US-09-938-719-5

; Sequence 5, Application US/09938719

; Patent No. US20020106742A1

; GENERAL INFORMATION:

; APPLICANT: SAMSON, MICHEL

; PARMENTIER, MARC

; VASSART, GILBERT

; LIBERT, FREDERICK

; TITLE OF INVENTION: ACTIVE AND INACTIVE CC-CHEMOKINES RECEPTOR

; AND NUCLEIC ACID MOLECULES ENCODING SAID RECEPTOR

; NUMBER OF SEQUENCES: 17

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Knobbe, Martens, Olson & Bear

; STREET: 620 Newport Center Drive 16th Floor

; CITY: Newport Beach

; STATE: CA

; COUNTRY: U.S.A.

; ZIP: 92660

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk

; COMPUTER: IBM PC compatible

; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: PatentIn Release #1.0, Version #1.25 (EPO)

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/09/938,719

; FILING DATE: 24-Aug-2001

; CLASSIFICATION: <Unknown>

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: 09/626,939

; FILING DATE: 27-JULY-2000

; ATTORNEY/AGENT INFORMATION:

; NAME: Altman, Daniel E

; REGISTRATION NUMBER: 34,115

```

;          REFERENCE/DOCKET NUMBER: <Unknown>
;  INFORMATION FOR SEQ ID NO: 5:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH: 352 amino acids
;      TYPE: amino acid
;      TOPOLOGY: linear
;    MOLECULE TYPE: protein
;    SEQUENCE DESCRIPTION: SEQ ID NO: 5:
US-09-938-719-5

```

```

Query Match          47.6%;  Score 379;  DB 9;  Length 352;
Best Local Similarity 35.0%;  Pred. No. 3e-29;
Matches 99;  Conservative 7;  Mismatches 27;  Indels 150;  Gaps 7;

```

```

Qy      1 MDYQVSSPIYDINYYTSEPCQKINVKQIAA----- 30
          ||||||||||||||||||||||||||||
Db      1 MDYQVSSPIYDINYYTSEPCQKINVKQIAARLLPPLYSLVFIFGFVGNMLVILILINCKR 60

Qy     31 -----YKCGLC-----AAAQWDFGNTMCQHQRVHGHHSYKCG--- 64
          |  |                      |||||||||  | : | :  |
Db     61 LKSMTDIYLLNLAISDLFFLLTVPFWAHYAAAQWDFGNTMC--QLLTGLYFIGFFSGIFF 118

Qy     65 -----LCTRSQKEGLHYTC 78
          : |||||||||
Db    119 IILLTIDRYLAVVHAVFALKARTVTFGVVTSVITWVAVFASLPGIIFTRSQKEGLHYTC 178

Qy     79 SSHFPYSQYQFWKNFQTLKI-----HQRVHGG----- 105
          ||||||||||||||||  :: |
Db    179 SSHFPYSQYQFWKNFQTLKIVILGLVLP LLVMVICYS GILKTL LRCRNEKKRHRAVRLIF 238

Qy    106 -----GGSYKCGLC---QEFFGLNNC SSSNRLDGHQRV 135
          |  |  |||||||||||||  :|
Db    239 TIMIVYFLFWAPYNIVLLLNTFQEFFGLNNC SSSNRLDQAMQV 281

```

RESULT 13

US-09-939-226-5

; Sequence 5, Application US/09939226

; Patent No. US20020110805A1

; GENERAL INFORMATION:

; APPLICANT: SAMSON, MICHEL

; PARMENTIER, MARC

; VASSART, GILBERT

; LIBERT, FREDERICK

; TITLE OF INVENTION: ACTIVE AND INACTIVE CC-CHEMOKINES RECEPTOR

; AND NUCLEIC ACID MOLECULES ENCODING SAID RECEPTOR

; NUMBER OF SEQUENCES: 17

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Knobbe, Martens, Olson & Bear

; STREET: 620 Newport Center Drive 16th Floor

; CITY: Newport Beach

; STATE: CA

; COUNTRY: U.S.A.

; ZIP: 92660

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk

; COMPUTER: IBM PC compatible

```

;      OPERATING SYSTEM: PC-DOS/MS-DOS
;      SOFTWARE: PatentIn Release #1.0, Version #1.25 (EPO)
;
;      CURRENT APPLICATION DATA:
;      APPLICATION NUMBER: US/09/939,226
;      FILING DATE: 24-Aug-2001
;      CLASSIFICATION: <Unknown>
;
;      PRIOR APPLICATION DATA:
;      APPLICATION NUMBER: 09/626,939
;      FILING DATE: 2000-07-27
;
;      ATTORNEY/AGENT INFORMATION:
;      NAME: Altman, Daniel E
;      REGISTRATION NUMBER: 34,115
;      REFERENCE/DOCKET NUMBER: <Unknown>
;
;      INFORMATION FOR SEQ ID NO: 5:
;      SEQUENCE CHARACTERISTICS:
;      LENGTH: 352 amino acids
;      TYPE: amino acid
;      TOPOLOGY: linear
;      MOLECULE TYPE: protein
;      SEQUENCE DESCRIPTION: SEQ ID NO: 5:
US-09-939-226-5

```

```

Query Match          47.6%;  Score 379;  DB 9;  Length 352;
Best Local Similarity 35.0%;  Pred. No. 3e-29;
Matches 99;  Conservative 7;  Mismatches 27;  Indels 150;  Gaps 7;

```

```

Qy      1 MDYQVSSPIYDINYYTSEPCQKINVKQIAA----- 30
          |||
Db      1 MDYQVSSPIYDINYYTSEPCQKINVKQIAARLLPPLYSLVFIFGFVGNMLVILILINCKR 60

Qy      31 -----YKCGLC-----AAAQWDFGNTMCQHQRVHGHHSYKCG--- 64
          |  |                      |||  | : | : |
Db      61 LKSMTDIYLLNLAISDLFFLLTVPFWAHYAAAQWDFGNTMC--QLLTGLYFIGFFSGIFF 118

Qy      65 -----LCTRSQKEGLHYTC 78
          : |||
Db      119 IILLTIDRYLAVVHAVFALKARTVTFGVVTSVITWVVAVFASLPGLIFTRSQKEGLHYTC 178

Qy      79 SSHFPYSQYQFWKNFQTLKI-----HQRVHGG----- 105
          |||                      :|
Db      179 SSHFPYSQYQFWKNFQTLKIVILGLVLP LLVMVICYS GILKTL LRCRNEKKRHRAVRLIF 238

Qy      106 -----GGSYKCGLC---QEFFGLNNC SSSNRLDGHQRV 135
          |  | |||  :|
Db      239 TIMIVYFLFWAPYNIVLLLNTFQEFFGLNNC SSSNRLDQAMQV 281

```

# RESULT 14

US-09-938-703-5

; Sequence 5, Application US/09938703

; Patent No. US20020110870A1

; GENERAL INFORMATION:

; APPLICANT: SAMSON, MICHEL

; PARMENTIER, MARC

; VASSART, GILBERT

; LIBERT, FREDERICK

; TITLE OF INVENTION: ACTIVE AND INACTIVE CC-CHEMOKINES RECEPTOR

```

;                               AND NUCLEIC ACID MOLECULES ENCODING SAID RECEPTOR
;
;   NUMBER OF SEQUENCES: 17
;   CORRESPONDENCE ADDRESS:
;       ADDRESSEE: Knobbe, Martens, Olson & Bear
;       STREET: 620 Newport Center Drive 16th Floor
;       CITY: Newport Beach
;       STATE: CA
;       COUNTRY: U.S.A.
;       ZIP: 92660
;   COMPUTER READABLE FORM:
;       MEDIUM TYPE: Floppy disk
;       COMPUTER: IBM PC compatible
;       OPERATING SYSTEM: PC-DOS/MS-DOS
;       SOFTWARE: PatentIn Release #1.0, Version #1.25 (EPO)
;   CURRENT APPLICATION DATA:
;       APPLICATION NUMBER: US/09/938,703
;       FILING DATE: 24-Aug-2001
;       CLASSIFICATION: <Unknown>
;   PRIOR APPLICATION DATA:
;       APPLICATION NUMBER: 09/626,939
;       FILING DATE: 2000-07-27
;   ATTORNEY/AGENT INFORMATION:
;       NAME: Altman, Daniel E
;       REGISTRATION NUMBER: 34,115
;       REFERENCE/DOCKET NUMBER: <Unknown>
;   INFORMATION FOR SEQ ID NO: 5:
;       SEQUENCE CHARACTERISTICS:
;           LENGTH: 352 amino acids
;           TYPE: amino acid
;           TOPOLOGY: linear
;       MOLECULE TYPE: protein
;       SEQUENCE DESCRIPTION: SEQ ID NO: 5:
US-09-938-703-5

```

```

Query Match          47.6%;  Score 379;  DB 9;  Length 352;
Best Local Similarity 35.0%;  Pred. No. 3e-29;
Matches 99;  Conservative 7;  Mismatches 27;  Indels 150;  Gaps 7;

```

```

Qy      1 MDYQVSSPIYDINYYTSEPCQKINVKQIAA----- 30
          |||||
Db      1 MDYQVSSPIYDINYYTSEPCQKINVKQIAARLLPPLYSLVFIFGFVGNMLVILILINCKR 60

Qy      31 -----YKCGLC-----AAAQWDFGNTMCQHQRVHGHHSYKCG--- 64
          |  |               |||||  | : | : |
Db      61 LKSMTDIYLLNLAISDLFFLLTVPEFWAHYAAAQWDFGNTMC--QLLTGLYFIGFFSGIFF 118

Qy      65 -----LCTRSQKEGLHYTC 78
          : |||||
Db      119 IILLTIDRYLAVVHAVFALKARTVTFGVVTSVITWVAVFASLPGIIFTRSQKEGLHYTC 178

Qy      79 SSHFPYSQYQFWKNFQTLKI-----HQRVHGG----- 105
          |||||  :: |
Db      179 SSHFPYSQYQFWKNFQTLKIVILGLVLPLLMVICYSGILKTLRLCRNEKKRHRVRLIF 238

Qy      106 -----GGSYKCGLC----QEFGGLNNCSSSNRLDGHQRV 135
          |  |               |||||  :|
Db      239 TIMIVYFLFWAPYNIVLLNLTQEFGLNNCSSSNRLDQAMQV 281

```

RESULT 15  
 US-09-502-783A-2  
 ; Sequence 2, Application US/09502783A  
 ; Patent No. US20020132269A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Li, Yi  
 ; APPLICANT: Ruben, Steven M.  
 ; TITLE OF INVENTION: Polynucleotides Encoding Human G-Protein Chemokine  
 Receptor (CCR5)  
 ; TITLE OF INVENTION: HDGNR10  
 ; FILE REFERENCE: 1488.1150006  
 ; CURRENT APPLICATION NUMBER: US/09/502,783A  
 ; CURRENT FILING DATE: 2001-08-23  
 ; PRIOR APPLICATION NUMBER: 08/466,343  
 ; PRIOR FILING DATE: 1995-06-06  
 ; NUMBER OF SEQ ID NOS: 9  
 ; SOFTWARE: PatentIn version 3.0  
 ; SEQ ID NO 2  
 ; LENGTH: 352  
 ; TYPE: PRT  
 ; ORGANISM: Homo sapiens  
 US-09-502-783A-2

Query Match 47.6%; Score 379; DB 9; Length 352;  
 Best Local Similarity 35.0%; Pred. No. 3e-29;  
 Matches 99; Conservative 7; Mismatches 27; Indels 150; Gaps 7;

```

Qy      1 MDYQVSSPIYDINYYTSEPCQKINVKQIAA----- 30
          |||
Db      1 MDYQVSSPIYDINYYTSEPCQKINVKQIAARLLPPLYSLVFIFGFVGNMLVILILINCKR 60

Qy      31 -----YKCGLC-----AAAQWDFGNTMCQHQRVHGHHHSYKCG--- 64
          |  |
Db      61 LKSMTDIYLLNLAISDLFFLLTVPFWAHYAAAQWDFGNTMC--QLLTGLYFIGFFSFIGFF 118

Qy      65 -----LCTRSQKEGLHYTC 78
          : |||
Db      119 IILLTIDRYLAVVHAVFALKARTVTFGVVTSVITWVAVFASLPGLIIFTRSQKEGLHYTC 178

Qy      79 SSHFPYSQYQFWKNFQTLKI-----HQRVHGG----- 105
          |||
Db      179 SSHFPYSQYQFWKNFQTLKIVILGLVLP LLVMVICYSGILKTLLRCRNEKKRHRAVRLIF 238

Qy      106 -----GGSYKCGLC-----QEFGGLNNCSSSNRLDGHQRV 135
          |  |
Db      239 TIMIVYFLWAPYNIVLLNLTQEFGLNNCSSSNRLDQAMQV 281
  
```

Search completed: September 17, 2004, 17:12:52  
 Job time : 86.7356 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: September 17, 2004, 16:42:06 ; Search time 66.8949 Seconds  
(without alignments)  
650.894 Million cell updates/sec

Title: US-10-057-890A-10  
Perfect score: 797  
Sequence: 1 MDYQVSSPIYDINYYTSEPC.....GLNNCSSSNRLDGHQRVHAA 138

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 1017041 seqs, 315518202 residues

Total number of hits satisfying chosen parameters: 1017041

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : SPTREMBL\_25:\*  
1: sp\_archaea:\*  
2: sp\_bacteria:\*  
3: sp\_fungi:\*  
4: sp\_human:\*  
5: sp\_invertebrate:\*  
6: sp\_mammal:\*  
7: sp\_mhc:\*  
8: sp\_organelle:\*  
9: sp\_phage:\*  
10: sp\_plant:\*  
11: sp\_rodent:\*  
12: sp\_virus:\*  
13: sp\_vertibrate:\*  
14: sp\_unclassified:\*  
15: sp\_rvirus:\*  
16: sp\_bacteriap:\*  
17: sp\_archeap:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result	Score	Query	Match	Length	DB	ID	Description
No.							

---





1	373	46.8	352	6	O18772	O18772	pan troglod
2	370	46.4	344	6	Q9TQR8	Q9tqr8	cercocobus
3	370	46.4	344	6	O77833	O77833	cercocobus
4	369	46.3	352	6	Q95NC5	Q95nc5	hylobates s
5	369	46.3	352	6	Q9TV50	Q9tv50	pan troglod
6	367	46.0	352	6	Q95NC3	Q95nc3	miopithecus
7	367	46.0	352	6	O18771	O18771	pan troglod
8	366	45.9	352	6	Q9TSK1	Q9tsk1	cercopithec
9	366	45.9	352	6	Q9TV49	Q9tv49	cercocobus
10	366	45.9	352	6	Q9XT13	Q9xt13	papio anubi
11	366	45.9	352	6	Q9TV45	Q9tv45	cercopithec
12	366	45.9	352	6	Q95NE8	Q95ne8	cercopithec
13	365	45.8	352	6	Q9XT14	Q9xt14	colobus gue
14	365	45.8	352	6	Q95NC6	Q95nc6	trachypithe
15	365	45.8	352	6	Q95NC8	Q95nc8	colobus pol
16	364	45.7	352	6	Q9XS99	Q9xs99	gorilla gor
17	363	45.5	352	6	O97975	O97975	macaca arct
18	363	45.5	352	6	Q9MZA2	Q9mza2	cercopithec
19	362	45.4	352	6	Q95NC1	Q95nc1	theropithec
20	362	45.4	352	6	Q95NC0	Q95nc0	hylobates m
21	362	45.4	352	6	Q9TV46	Q9tv46	cercopithec
22	362	45.4	352	6	O97962	O97962	pygathrix a
23	361	45.3	352	6	Q95NC7	Q95nc7	nasalis lar
24	361	45.3	352	6	Q9TV42	Q9tv42	cercopithec
25	361	45.3	352	6	O77776	O77776	cercocobus
26	361	45.3	352	6	Q95NE1	Q95ne1	cercocobus
27	361	45.3	352	6	Q9TQX0	Q9tqx0	cercopithec
28	359	45.0	352	6	Q9TV43	Q9tv43	cercopithec
29	358	44.9	352	6	O18770	O18770	pan troglod
30	357	44.8	352	6	Q9TV47	Q9tv47	cercopithec
31	357	44.8	352	6	Q9BGN5	Q9bgn5	cercopithec
32	357	44.8	352	6	Q9XT12	Q9xt12	cercopithec
33	357	44.8	352	6	Q9TSQ7	Q9tsq7	cercopithec
34	356	44.7	352	6	Q95ND2	Q95nd2	mandrillus
35	356	44.7	352	6	Q95ND1	Q95nd1	mandrillus
36	355	44.5	352	6	Q9XT76	Q9xt76	cercopithec
37	355	44.5	352	6	Q9TV44	Q9tv44	cercopithec
38	355	44.5	352	6	Q9XS35	Q9xs35	macaca neme
39	355	44.5	352	6	Q9MZA3	Q9mza3	hylobates a
40	355	44.5	352	6	Q9TV93	Q9tv93	macaca arct
41	355	44.5	352	6	Q95ND0	Q95nd0	erythrocebu
42	346	43.4	352	6	Q9BGN6	Q9bgn6	cercopithec
43	345	43.3	339	4	Q9UN25	Q9un25	homo sapien
44	344	43.2	339	4	Q9UN23	Q9un23	homo sapien
45	344	43.2	339	4	Q9UN27	Q9un27	homo sapien

# ALIGNMENTS

## RESULT 1

O18772

ID O18772 PRELIMINARY; PRT; 352 AA.

AC O18772;

DT 01-JAN-1998 (TrEMBLrel. 05, Created)

DT 01-JAN-1998 (TrEMBLrel. 05, Last sequence update)

DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)

DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)  
 DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)  
 DE Chemokine receptor CCR5.  
 GN CCR5.  
 OS Cercopithecus torquatus torquatus, and  
 OS Cercopithecus torquatus atys (Red-crowned mangabey) (Sooty mangabey).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Cercopithecidae;  
 OC Cercopithecinae; Cercopithecus.  
 OX NCBI\_TaxID=81944, 9531;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC SPECIES=C.torquatus torquatus; STRAIN=1049, and 997;  
 RA Chen Z., Kwon D., Jin Z., Monard S., Telfer P., Jones M., Lu C.,  
 RA Aguilar R., Ho D.D., Marx P.A.;  
 RT "Natural infection of a homozygous delta 24 CCR5 red-capped mangabey  
 RT with a R2b-tropic simian immunodeficiency virus.";  
 RL J. Exp. Med. 0:0-0(1998).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RC SPECIES=C.torquatus atys;  
 RA Palacios E., Digilio L., McClure H.M., Chen Z., Marx P.A.,  
 RA Goldsmith M.A., Grant R.M.;  
 RT "Parallel evolution of CCR5-null phenotypes in humans and in a natural  
 RT host of simian immunodeficiency virus.";  
 RL Curr. Biol. 0:0-0(1998).  
 RN [3]  
 RP SEQUENCE FROM N.A.  
 RC SPECIES=C.torquatus torquatus; STRAIN=RCM411;  
 RA Beer B.E., Kuiken C.L., Tooze Z., Foley B.T., Goeken R.M., Brown C.R.,  
 RA St Claire M., Hirsch V.M.;  
 RT "Characterization of novel simian immunodeficiency viruses from  
 RT redcapped mangabeys from Nigeria (SIVrcmNgM411 and NgD409).";  
 RL Submitted (FEB-2001) to the EMBL/GenBank/DDBJ databases.  
 DR EMBL; AF094753; AAC62474.1; -.  
 DR EMBL; AF079473; AAC31194.1; -.  
 DR EMBL; AF084003; AAC62471.1; -.  
 DR EMBL; AF349683; AAK69685.1; -.  
 DR GO; GO:0016021; C:integral to membrane; IEA.  
 DR GO; GO:0004872; F:receptor activity; IEA.  
 DR GO; GO:0001584; F:rhodopsin-like receptor activity; IEA.  
 DR GO; GO:0007186; P:G-protein coupled receptor protein signalin. . .; IEA.  
 DR InterPro; IPR000276; GPCR\_Rhodpsn.  
 DR Pfam; PF00001; 7tm\_1; 1.  
 DR PRINTS; PR00237; GPCRRHODOPSN.  
 DR PROSITE; PS00237; G\_PROTEIN\_RECEP\_F1\_1; 1.  
 DR PROSITE; PS50262; G\_PROTEIN\_RECEP\_F1\_2; 1.  
 KW Receptor.  
 SQ SEQUENCE 344 AA; 39592 MW; E15F5F601191A4D1 CRC64;

Query Match 46.4%; Score 370; DB 6; Length 344;  
 Best Local Similarity 34.9%; Pred. No. 4e-33;  
 Matches 96; Conservative 9; Mismatches 28; Indels 142; Gaps 7;

Qy 1 MDYQVSSPIYDINYTSEPCQKINVKQIAA----- 30  
 ||||| |||:|||||||  
 Db 1 MDYQVSSPTYDIDYYTSEPCQKINVKQIAARLLPPLYSLVFIFGFVGNILVVLILINCKR 60

```

Qy      31 -----YKCGLC-----AAAQWDFGNTMCQHQRVHGHHHSYKCG--- 64
          | |               ||||| | | : | : : |
Db      61 LKSMTDIYLLNLAISDLLFLLTVPFWAHYAAAQWDFGNTMC--QLLTGLYFIGFFSGIFF 118

Qy      65 -----LCTRSQKEGLHYTCSSHPYSQ 86
          : ||||:|||||
Db     119 IILLTIDRYLAIVHAVFALKARTVTFGVVAVFASLPGIIFTRSQREGLHYTCSSHPYSQ 178

Qy      87 YQFWKNFQTLKI-----HQRVHGG----- 105
          ||||| | | : |
Db     179 YQFWKNFQTLKIVILGLVLPLLVMVICYSGILKTLRLCRNEKKRRAVRLIFTIMIVYFL 238

Qy     106 -GGSYKCGLC----QEFFGLNNCSSSNRLDGHQRV 135
          | | | ||||| | | : |
Db     239 FWAPYNIVLLLNTFQEFFGLNNCSSSNRLDQAMQV 273

```

RESULT 3

O77833

```

ID   O77833      PRELIMINARY;      PRT;   344 AA.
AC   O77833;
DT   01-NOV-1998 (TrEMBLrel. 08, Created)
DT   01-NOV-1998 (TrEMBLrel. 08, Last sequence update)
DT   01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE   Chemokine receptor CCR5.
GN   CCR5.
OS   Cercopithecus torquatus torquatus.
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Cercopithecidae;
OC   Cercopithecinae; Cercopithecus.
OX   NCBI_TaxID=81944;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=1208, and 009;
RA   Chen Z., Kwon D., Jin Z., Monard S., Telfer P., Jones M., Lu C.,
RA   Aguilar R., Ho D.D., Marx P.A.;
RT   "Natural infection of a homozygous delta 24 CCR5 red-capped mangabey
RT   with a R2b-tropic simian immunodeficiency virus.";
RL   J. Exp. Med. 0:0-0(1998).
DR   EMBL; AF094752; AAC62473.1; -.
DR   EMBL; AF084002; AAC62470.1; -.
DR   GO; GO:0016021; C:integral to membrane; IEA.
DR   GO; GO:0001584; F:rhodopsin-like receptor activity; IEA.
DR   GO; GO:0007186; P:G-protein coupled receptor protein signalin. . .; IEA.
DR   InterPro; IPR000276; GPCR_Rhodpsn.
DR   Pfam; PF00001; 7tm_1; 1.
DR   PRINTS; PR00237; GPCR_RHODOPSN.
DR   PROSITE; PS00237; G_PROTEIN_RECEP_F1_1; 1.
DR   PROSITE; PS50262; G_PROTEIN_RECEP_F1_2; 1.
SQ   SEQUENCE   344 AA;  39578 MW;  0CEC05E47C2F6DE6 CRC64;

```

```

Query Match          46.4%;  Score 370;  DB 6;  Length 344;
Best Local Similarity 34.9%;  Pred. No. 4e-33;
Matches 96;  Conservative 9;  Mismatches 28;  Indels 142;  Gaps 7;

```

```

Qy      1 MDYQVSSPIYDINYYTSEPCQKINVKQIAA----- 30

```

```

      ||||| |||:|||||
Db      1 MDYQVSSPTYDIDYYTSEPCQKINVKQIAARLLPPLYSLVFVFGFVGNILVVLILINCKR 60
Qy      31 -----YKCGLC-----AAAQWDFGNTMCQHQRVHGHHSYKCG--- 64
      | | ||||| | : | : |
Db      61 LKSMTDIYLLNLAIISDLLFLLTVPFWAHYAAAQWDFGNTMC--QLLTGLYFIGFFSGIFF 118
Qy      65 -----LCTRSQKEGLHYTCSSSHFPYSQ 86
      : ||||:|||||
Db      119 IILLTIDRYLAIVHAVFALKARTVTFGVVAVFASLPGIIFTRSQREGLHYTCSSSHFPYSQ 178
Qy      87 YQFWKNFQTLKI-----HQRVHGG----- 105
      ||||| | : |
Db      179 YQFWKNFQTLKIVILGLVLP LLVMVICYSGILKTLRLCRNEKKRHRAVRLIFTIMIVYFL 238
Qy      106 -GGSYKCGLC----QEFFGLNNC SSSNRLDGHQRV 135
      | | ||||| | : |
Db      239 FWAPYNIVLLLNTFQEFFGLNNC SSSNRLDQAMQV 273

```

#### RESULT 4

##### Q95NC5

```

ID   Q95NC5          PRELIMINARY;      PRT;   352 AA.
AC   Q95NC5;
DT   01-DEC-2001 (TrEMBLrel. 19, Created)
DT   01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT   01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE   C-C chemokine receptor 5.
GN   CCR5.
OS   Hylobates syndactylus (Siamang) (Symphalangus syndactylus).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hylobatidae; Hylobates.
OX   NCBI_TaxID=9590;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Zhang Y., Ryder O.A., Zhang Y.;
RT   "Sequence comparison of the CCR5 gene in primates and primate
RT   phylogeny.";
RL   Submitted (AUG-1999) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AF177884; AAK43367.1; -.
DR   GO; GO:0016021; C:integral to membrane; IEA.
DR   GO; GO:0004872; F:receptor activity; IEA.
DR   GO; GO:0001584; F:rhodopsin-like receptor activity; IEA.
DR   GO; GO:0007186; P:G-protein coupled receptor protein signalin. . .; IEA.
DR   InterPro; IPR000276; GPCR_Rhodpsn.
DR   Pfam; PF00001; 7tm_1; 1.
DR   PROSITE; PS00237; G_PROTEIN_RECEP_F1_1; 1.
DR   PROSITE; PS50262; G_PROTEIN_RECEP_F1_2; 1.
KW   Receptor.
SQ   SEQUENCE   352 AA;  40508 MW;  F4F64B3AD5AF658A CRC64;

```

```

Query Match          46.3%;  Score 369;  DB 6;  Length 352;
Best Local Similarity 34.3%;  Pred. No. 5.3e-33;
Matches   97;  Conservative    8;  Mismatches   28;  Indels   150;  Gaps    7;

```

```

Qy      1 MDYQVSSPIYDINYYTSEPCQKINVKQIAA----- 30
      ||||| |||:|||||

```

```

Db          1 MDYQVSSPTYDIDYYTSEPCQKINVKQIAARLLPPLYSLVFIFGFVGNMLVILILINCKR 60
Qy          31 -----YKCGLC-----AAAQWDFGNTMCQHQRVHGHHSYKCG--- 64
              | |               ||||| | : | : |
Db          61 LKSMTDIYLLNLAISDLFFLLTVPFWAHYAAAQWDFGNTMC--QLLTGLYFIGFFSGIFF 118
Qy          65 -----LCTRSQKEGLHYTC 78
                          : |||||
Db          119 IILLTIDRYLAIVHAVFALKARTVTFGVTVSVITWVAVFASLPGIIFTRSQKEGLHYTC 178
Qy          79 SSHFPYSQYQFWKNFQTLKI-----HQRVHGG----- 105
              ||||| | : |
Db          179 SSHFPYSQYQFWKNFQTLKIVILGLVLPLLVMVICYSGILKTLRLCRNEKKRHRVRLIF 238
Qy          106 -----GGSYKCGLC---QEFGFLNNCSSSNRLDGHQRV 135
              | |               ||||| | : |
Db          239 TIMIVYFLFWAPYNIVLLNTFQEFFGLNNCSSSNRLDQAMQV 281

```

RESULT 5

Q9TV50

```

ID   Q9TV50          PRELIMINARY;          PRT;    352 AA.
AC   Q9TV50;
DT   01-MAY-2000 (TrEMBLrel. 13, Created)
DT   01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT   01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE   CC chemokine receptor type 5.
GN   CCR5.
OS   Pan troglodytes (Chimpanzee).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Pan.
OX   NCBI_TaxID=9598;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=1410;
RX   MEDLINE=99335215; PubMed=10408730;
RA   Muller-Trutwin M.C., Corbet S., Hansen J., Georges-Courbot M.C.,
RA   Diop O., Rigoulet J., Barre-Sinoussi F., Fomsgaard A.;
RT   "Mutations in CCR5-coding sequences are not associated with SIV
RT   carrier status in African nonhuman primates.";
RL   AIDS Res. Hum. Retroviruses 15:931-939(1999).
RN   [2]
RP   SEQUENCE FROM N.A.
RC   STRAIN=1410;
RA   Mueller-Trutwin M.C., Corbet S., Hansen J., Diop O.,
RA   Georges-Courbot M.-C., Barre-Sinoussi F., Fomsgaard A.;
RL   Submitted (NOV-1997) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AF035214; AAD44007.1; -.
DR   GO; GO:0016021; C:integral to membrane; IEA.
DR   GO; GO:0004872; F:receptor activity; IEA.
DR   GO; GO:0001584; F:rhodopsin-like receptor activity; IEA.
DR   GO; GO:0007186; P:G-protein coupled receptor protein signalin. . .; IEA.
DR   InterPro; IPR000276; GPCR_Rhodpsn.
DR   Pfam; PF00001; 7tm_1; 1.
DR   PRINTS; PR00237; GPCRRHODOPSN.
DR   PROSITE; PS00237; G_PROTEIN_RECEP_F1_1; 1.
DR   PROSITE; PS50262; G_PROTEIN_RECEP_F1_2; 1.

```

KW Receptor.

SQ SEQUENCE 352 AA; 40481 MW; 2578A0F2C07A4A65 CRC64;

Query Match 46.3%; Score 369; DB 6; Length 352;

Best Local Similarity 34.3%; Pred. No. 5.3e-33;

Matches 97; Conservative 8; Mismatches 28; Indels 150; Gaps 7;

```
Qy      1 MDYQVSSPIYDINYYTSEPCQKINVKQIAA----- 30
          ||||| |||:|||||
Db      1 MDYQVSSPTYDIDYYTSEPCQKINVKQIAARLLPPLYSLVFIFGFVGNILVVLILINCKR 60

Qy     31 -----YKCGLC-----AAAQWDFGNTMCQHQRVHGHHSYKCG--- 64
          | | ||||| | : | : |
Db     61 LKSMTDIYLLNLAISDLFFLLTVPFWAHYAAAQWDFGNTMC--QLLTGLYFIGFFSGIFF 118

Qy     65 -----LCTRSQKEGLHYTC 78
          : |||||
Db    119 IILLSIDRYLAIVHAVFALKARTVTFGVVTSVITWVAVFASLPGIIFTRSQKEGLHYTC 178

Qy     79 SSHFPYSQYQFWKNFQTLKI-----HQRVHGG----- 105
          ||||| ||||| : |
Db    179 SSHFPYSQYQFWKNFQTLKIVILGLVLPLLMVICYSGILKTLRLCRNEKKRHRVRLIF 238

Qy    106 -----GGSYKCGLC----QEFFGLNNCSSSNRLDGHQRV 135
          | | ||||| : |
Db    239 TIMIVYFLFWAPYNIVLLLNTFQEFFGLNNCSSSNRLDQAMQV 281
```

#### RESULT 6

Q95NC3

ID Q95NC3 PRELIMINARY; PRT; 352 AA.

AC Q95NC3;

DT 01-DEC-2001 (TrEMBLrel. 19, Created)

DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)

DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)

DE C-C chemokine receptor 5.

GN CCR5.

OS Miopithecus talapoin (Talapoin) (Cercopithecus talapoin).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Primates; Catarrhini; Cercopithecidae;

OC Cercopithecinae; Miopithecus.

OX NCBI\_TaxID=36231;

RN [1]

RP SEQUENCE FROM N.A.

RA Zhang Y., Ryder O.A., Zhang Y.;

RT "Sequence comparison of the CCR5 gene in primates and primate  
phylogeny.";

RL Submitted (AUG-1999) to the EMBL/GenBank/DDBJ databases.

DR EMBL; AF177886; AAK43369.1; -.

DR GO; GO:0016021; C:integral to membrane; IEA.

DR GO; GO:0004872; F:receptor activity; IEA.

DR GO; GO:0001584; F:rhodopsin-like receptor activity; IEA.

DR GO; GO:0007186; P:G-protein coupled receptor protein signalin. . .; IEA.

DR InterPro; IPR000276; GPCR\_Rhodpsn.

DR Pfam; PF00001; 7tm\_1; 1.

DR PRINTS; PR00237; GPCRRHODOPSN.

DR PROSITE; PS00237; G\_PROTEIN\_RECEP\_F1\_1; 1.

DR PROSITE; PS50262; G\_PROTEIN\_RECEP\_F1\_2; 1.

KW Receptor..

SQ SEQUENCE 352 AA; 40546 MW; 6464152F3E566AE5 CRC64;

Query Match 46.0%; Score 367; DB 6; Length 352;

Best Local Similarity 33.9%; Pred. No. 8.9e-33;

Matches 96; Conservative 9; Mismatches 28; Indels 150; Gaps 7;

```
Qy      1 MDYQVSSPIYDINYYTSEPCQKINVKQIAA----- 30
          ||||| ||||| ||||| ||||| |||||
Db      1 MDYQVSSPTYDINYYTSEPCQKINVKQIAARLLPPLYSLVFIFGFVGNILVVLILINCKR 60

Qy     31 -----YKCGLC-----AAAQWDFGNTMCQHQRVHGHHHSYKCG--- 64
          |  |                               ||||| : : | : |
Db     61 LKSMTDIYLLNLAISDLLFLLTVPFWAHYAAAQWDFGNTMC--RLLTGLYFIGFFSGIFF 118

Qy     65 -----LCTRSQKEGLHYTC 78
          : |||: |||||
Db    119 IILLTIDRYLAIVHAVFALKARTVTFGVVTSVITWVAVFASLPGIIFTRSQREGLHYTC 178

Qy     79 SSHFPYSQYQFWKNFQTLKI-----HQRVHGG----- 105
          ||||| ||||| ||||| : : |
Db    179 SSHFPYSQYQFWKNFQTLKIVILGLVPLLVVICYSILKTLRLCRNEKKRHRVRLIF 238

Qy    106 -----GGSYKCGLC----QEFFGLNNCSSSNRLDGHQRV 135
          |  | ||||| :|
Db    239 TIMIVYFLFWAPYNIVLLNTFQEFFGLNNCSSSNRLDQAMQV 281
```

#### RESULT 7

O18771

ID O18771 PRELIMINARY; PRT; 352 AA.

AC O18771;

DT 01-JAN-1998 (TrEMBLrel. 05, Created)

DT 01-JAN-1998 (TrEMBLrel. 05, Last sequence update)

DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)

DE CCR5 receptor (Fragment).

GN CCR5.

OS Pan troglodytes (Chimpanzee).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Pan.

OX NCBI\_TaxID=9598;

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN=ChCCR5-141a;

RA Zhang L., Carruthers C.D., He T., Huang Y., Cao Y., Wang G., Hahn B.,

RA Ho D.D.;

RT "HIV-1 subtypes, co-receptor usage, and CCR5 polymorphism.";

RL AIDS Res. Hum. Retroviruses 0:0-0(1997).

DR EMBL; AF011539; AAB65739.1; -.

DR GO; GO:0016021; C:integral to membrane; IEA.

DR GO; GO:0004872; F:receptor activity; IEA.

DR GO; GO:0001584; F:rhodopsin-like receptor activity; IEA.

DR GO; GO:0007186; P:G-protein coupled receptor protein signalin. . .; IEA.

DR InterPro; IPR000276; GPCR\_Rhodpsn.

DR Pfam; PF00001; 7tm\_1; 1.

DR PRINTS; PR00237; GPCRRHODOPSN.



DR PROSITE; PS00237; G\_PROTEIN\_RECEP\_F1\_1; 1.  
 DR PROSITE; PS50262; G\_PROTEIN\_RECEP\_F1\_2; 1.  
 KW Receptor.  
 FT NON\_TER 352 352  
 SQ SEQUENCE 352 AA; 40466 MW; 3FFFAC7ABAE1D4FB CRC64;

Query Match 46.0%; Score 367; DB 6; Length 352;  
 Best Local Similarity 34.3%; Pred. No. 8.9e-33;  
 Matches 97; Conservative 8; Mismatches 28; Indels 150; Gaps 7;

```

Qy      1 MDYQVSSPIYDINYYTSEPCQKINVKQIAA----- 30
          |||||:|||||
Db      1 MDYQVSSPIYDIDYYTSEPCQKINVKQIAARLLPPLYSILVFIFGFVGNMLVILILINCKR 60

Qy      31 -----YKCGLC-----AAAQWDFGNTMCQHQRVHGHHHSYKCG--- 64
          | | ||||| | : | : |
Db      61 LKSMTDIYLLNLAISNLFLLTVPFWAHYAAAQWDFGNTMC--QLLTGLYFIGFFSGIFF 118

Qy      65 -----LCTRSQKEGLHYTC 78
          : |||||
Db      119 IILLTIDRYLAIVHAVFALKARTVTFGVVTSVITWVAVFASLPGIIFTRSQKGGGLHYTC 178

Qy      79 SSHFPYSQYQFWKNFQTLKI-----HQRVHGG----- 105
          ||||| : |
Db      179 SSHFPYSQYQFWKNFQTLKIVILGLVLP LLVMVICYSGILKTLRLCRNEKKRHRAVRLIF 238

Qy      106 -----GGSYKCGLC-----QEFFGLNNC SSSNRLDGHQRV 135
          | | ||||| : |
Db      239 TIMIVYFLEWAPYNIVLLLNTFQEFFGLNNC SSSNRLDQAMQV 281
  
```

# RESULT 8

## Q9TSK1

ID Q9TSK1 PRELIMINARY; PRT; 352 AA.  
 AC Q9TSK1;  
 DT 01-MAY-2000 (TrEMBLrel. 13, Created)  
 DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)  
 DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)  
 DE G-protein coupled chemokine receptor.  
 OS Cercopithecus aethiops (Green monkey) (Grivet).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Cercopithecidae;  
 OC Cercopithecinae; Cercopithecus.  
 OX NCBI\_TaxID=9534;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA Holtkamp N., Baier M., Werner A.;  
 RL Submitted (AUG-1997) to the EMBL/GenBank/DDBJ databases.  
 DR EMBL; AF019379; AAD01639.1; -.  
 DR GO; GO:0016021; C:integral to membrane; IEA.  
 DR GO; GO:0004872; F:receptor activity; IEA.  
 DR GO; GO:0001584; F:rhodopsin-like receptor activity; IEA.  
 DR GO; GO:0007186; P:G-protein coupled receptor protein signalin. . .; IEA.  
 DR InterPro; IPR000276; GPCR\_Rhodpsn.  
 DR Pfam; PF00001; 7tm\_1; 1.  
 DR PRINTS; PR00237; GPCRRHODOPSN.  
 DR PROSITE; PS00237; G\_PROTEIN\_RECEP\_F1\_1; 1.

DR PROSITE; PS50262; G\_PROTEIN\_RECEP\_F1\_2; 1.  
KW Receptor.  
SQ SEQUENCE 352 AA; 40588 MW; 0F1869D9A6668DBB CRC64;

Query Match 45.9%; Score 366; DB 6; Length 352;  
Best Local Similarity 33.9%; Pred. No. 1.2e-32;  
Matches 96; Conservative 9; Mismatches 28; Indels 150; Gaps 7;

```
Qy      1 MDYQVSSPIYDINYYTSEPCQKINVKQIAA----- 30
          ||||| |||:|||||
Db      1 MDYQVSSPTYDIDYYTSEPCQKINVKQIAARLLPPLYSLVFIFGFVGNILVVLILINCKR 60

Qy     31 -----YKCGLC-----AAAQWDFGNTMCQHQRVHGHHHSYKCG--- 64
          |  |               ||||| | : | : |
Db     61 LKSMTDIYLLNLAISDLLFLLTVPFWAHYAAAQWDFGNTMC--QLLTGLYFIGFFSGIFF 118

Qy     65 -----LCTRSQKEGLHYTC 78
          : |||||
Db    119 IILLTIDRYLAIVHAVFALKARTVTFGVVTSVITWVAVFASLPRIIFTRSQREGLHYTC 178

Qy     79 SSHFPYSQYQFWKNFQTLKI-----HQRVHGG----- 105
          ||||| |||||
Db    179 SSHFPYSQYQFWKNFQTLKIVILGLVPLLVVICYSKILKTLRLCRNEKKRHRVRLIF 238

Qy    106 -----GGSYKCGLC----QEFFGLNNCSSSNRLDGHQRV 135
          |  | |||||
Db    239 TIMIVYFLFWAPYNIVLLLNLTQEFFGLNNCSSSNRLDQAMQV 281
```

#### RESULT 9

Q9TV49

ID Q9TV49 PRELIMINARY; PRT; 352 AA.  
AC Q9TV49;  
DT 01-MAY-2000 (TrEMBLrel. 13, Created)  
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)  
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)  
DE CC chemokine receptor type 5 (C-C chemokine receptor 5).  
GN CCR5.  
OS Cercopithecus galeritus (Agile mangabey).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Primates; Catarrhini; Cercopithecidae;  
OC Cercopithecinae; Cercopithecus.  
OX NCBI\_TaxID=9532;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=4;  
RX MEDLINE=99335215; PubMed=10408730;  
RA Muller-Trutwin M.C., Corbet S., Hansen J., Georges-Courbot M.C.,  
RA Diop O., Rigoulet J., Barre-Sinoussi F., Fomsgaard A.;  
RT "Mutations in CCR5-coding sequences are not associated with SIV  
RT carrier status in African nonhuman primates."  
RL AIDS Res. Hum. Retroviruses 15:931-939(1999).  
RN [2]  
RP SEQUENCE FROM N.A.  
RC STRAIN=4;  
RA Mueller-Trutwin M.C., Corbet S., Hansen J., Diop O.,  
RA Georges-Courbot M.-C., Barre-Sinoussi F., Fomsgaard A.;

```

RL Submitted (NOV-1997) to the EMBL/GenBank/DDBJ databases.
RN [3]
RP SEQUENCE FROM N.A.
RA Zhang Y., Ryder O.A., Zhang Y.;
RT "Sequence comparison of the CCR5 gene in primates and primate
RT phylogeny.";
RL Submitted (AUG-1999) to the EMBL/GenBank/DDBJ databases.
DR EMBL; AF035215; AAD44008.1; -.
DR EMBL; AF177898; AAK43381.1; -.
DR GO; GO:0016021; C:integral to membrane; IEA.
DR GO; GO:0004872; F:receptor activity; IEA.
DR GO; GO:0001584; F:rhodopsin-like receptor activity; IEA.
DR GO; GO:0007186; P:G-protein coupled receptor protein signalin. . .; IEA.
DR InterPro; IPR000276; GPCR_Rhodpsn.
DR Pfam; PF00001; 7tm_1; 1.
DR PRINTS; PR00237; GPCRRHODOPSN.
DR PROSITE; PS00237; G_PROTEIN_RECEP_F1_1; 1.
DR PROSITE; PS50262; G_PROTEIN_RECEP_F1_2; 1.
KW Receptor.
SQ SEQUENCE 352 AA; 40479 MW; 5E1504A9B87278B2 CRC64;

```

## RESULT 10



OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Cercopithecidae;  
 OC Cercopithecinae; Cercopithecus.  
 OX NCBI\_TaxID=36228;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=7;  
 RX MEDLINE=99335215; PubMed=10408730;  
 RA Muller-Trutwin M.C., Corbet S., Hansen J., Georges-Courbot M.C.,  
 RA Diop O., Rigoulet J., Barre-Sinoussi F., Fomsgaard A.;  
 RT "Mutations in CCR5-coding sequences are not associated with SIV  
 RT carrier status in African nonhuman primates.";  
 RL AIDS Res. Hum. Retroviruses 15:931-939(1999).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=7;  
 RA Mueller-Trutwin M.C., Corbet S., Hansen J., Diop O.,  
 RA Georges-Courbot M.-C., Barre-Sinoussi F., Fomsgaard A.;  
 RL Submitted (NOV-1997) to the EMBL/GenBank/DDBJ databases.  
 DR EMBL; AF035219; AAD44012.1; -.  
 DR GO; GO:0016021; C:integral to membrane; IEA.  
 DR GO; GO:0004872; F:receptor activity; IEA.  
 DR GO; GO:0001584; F:rhodopsin-like receptor activity; IEA.  
 DR GO; GO:0007186; P:G-protein coupled receptor protein signalin. . .; IEA.  
 DR InterPro; IPR000276; GPCR\_Rhodpsn.  
 DR Pfam; PF00001; 7tm\_1; 1.  
 DR PRINTS; PR00237; GPCRRHODOPSN.  
 DR PROSITE; PS00237; G\_PROTEIN\_RECEP\_F1\_1; 1.  
 DR PROSITE; PS50262; G\_PROTEIN\_RECEP\_F1\_2; 1.  
 KW Receptor.  
 SQ SEQUENCE 352 AA; 40658 MW; BC665891ED5661F8 CRC64;

Query Match 45.9%; Score 366; DB 6; Length 352;  
 Best Local Similarity 33.9%; Pred. No. 1.2e-32;  
 Matches 96; Conservative 9; Mismatches 28; Indels 150; Gaps 7;

QY 1 MDYQVSSPIYDINYTSEPCQKINVKQIAA----- 30  
 ||||| |||:|||||||  
 Db 1 MDYQVSSPTYDIDYTTSEPCQKINVKQIAARLLPLLYSLVFIFGFVGNILVVLILINCKR 60  
 QY 31 -----YKCGLC-----AAAQWDFGNTMCQHQRVHGHHSYKCG--- 64  
 | | ||||| | : | : |  
 Db 61 LKSMTDIYLLNLAISDLLFLLTVPFWAHYAAAQWDFGNTMC--QLLTGLYFIGFFSGIFF 118  
 QY 65 -----LCTRSQKEGLHYTC 78  
 : |||:|||||  
 Db 119 IILLTIDRYLAIVHAVFALKARTVTFGVVTSVITWVAVFASLPRIIFTRSQREGLHYTC 178  
 QY 79 SSHFPYSQYQFWKNEQTLKI-----HQRVHGG----- 105  
 ||||| ||||| : : |  
 Db 179 SSHFPYSQYQFWKNEQTLKIVILGLVLP LLVMVICYSGLKTLRLCRNEKKRHRVAMLIF 238  
 QY 106 -----GGSYKCGLC-----QEFFGLNNCSSSNRLDGHQRV 135  
 | | ||||| ||||| : |  
 Db 239 TIMIVYFLFWAPYNIVLLLNTFQEFFGLNNCSSSNRLDQAMQV 281

# RESULT 12

Q95NE8

ID Q95NE8 PRELIMINARY; PRT; 352 AA.  
AC Q95NE8;  
DT 01-DEC-2001 (TrEMBLrel. 19, Created)  
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)  
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)  
DE CC chemokine receptor type 5.  
GN CCR5.  
OS Cercopithecus tantalus.  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Primates; Catarrhini; Cercopithecidae;  
OC Cercopithecinae; Cercopithecus.  
OX NCBI\_TaxID=60712;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=99335215; PubMed=10408730;  
RA Muller-Trutwin M.C., Corbet S., Hansen J., Georges-Courbot M.C.,  
RA Diop O., Rigoulet J., Barre-Sinoussi F., Fomsgaard A.;  
RT "Mutations in CCR5-coding sequences are not associated with SIV  
RT carrier status in African nonhuman primates.";  
RL AIDS Res. Hum. Retroviruses 15:931-939(1999).  
DR EMBL; AF081577; AAD45495.1; -.  
DR GO; GO:0016021; C:integral to membrane; IEA.  
DR GO; GO:0001584; F:rhodopsin-like receptor activity; IEA.  
DR GO; GO:0007186; P:G-protein coupled receptor protein signalin. . .; IEA.  
DR InterPro; IPR000276; GPCR\_Rhodpsn.  
DR Pfam; PF00001; 7tm\_1; 1.  
DR PRINTS; PR00237; GPCRRHODOPSN.  
DR PROSITE; PS00237; G\_PROTEIN\_RECEP\_F1\_1; 1.  
DR PROSITE; PS50262; G\_PROTEIN\_RECEP\_F1\_2; 1.  
SQ SEQUENCE 352 AA; 40576 MW; 0F1869D9B7968DBB CRC64;

Query Match 45.9%; Score 366; DB 6; Length 352;  
Best Local Similarity 33.9%; Pred. No. 1.2e-32;  
Matches 96; Conservative 9; Mismatches 28; Indels 150; Gaps 7;

Qy 1 MDYQVSSPIYDINYYTSEPCQKINVKQIAA----- 30  
| | | | | | | | | | : | | | | | | | | | | |  
Db 1 MDYQVSSPTYDIDYYTSEPCQKINVKQIAARLLPPLYSLVFIFGFVGNILVVLILINCKR 60  
Qy 31 -----YKCGLC-----AAAQWDFGNTMCQHQRVHGHGHHHSYKCG--- 64  
| | | | | | | | | | | | | | | | : | : : |  
Db 61 LKSMTDIYLLNLAISDLLFLLTVPFWAHYAAAQWDFGNTMC--QLLTGLYFIGFFSGIFF 118  
Qy 65 -----LCTRSQKEGLHYTC 78  
: | | | : | | | | |  
Db 119 IILLTIDRYLAIVHAVFALKARTVTFGVVTSVITWVAVFASLPRIIFTRSQREGLHYTC 178  
Qy 79 SSHFPYSQYQFWKNFQTLKI-----HQRVHGG----- 105  
| | | | | | | | | | | | | | | | : : |  
Db 179 SSHFPYSQYQFWKNFQTLKIVILGLVLPPLVMVICYSKILKTLRCRNEKKRHRAVRLIF 238  
Qy 106 -----GGSYKCGLC----QEFFGLNNCSSSNRLDGHQRV 135  
| | | | | | | | | | | | | | : |  
Db 239 TIMIVYFLEWAPYNIVLLNLTQEFFGLNNCSSSNRLDQAMQV 281

RESULT 13

Q9XT14

ID Q9XT14 PRELIMINARY; PRT; 352 AA.  
 AC Q9XT14;  
 DT 01-NOV-1999 (TrEMBLrel. 12, Created)  
 DT 01-NOV-1999 (TrEMBLrel. 12, Last sequence update)  
 DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)  
 DE CC chemokine receptor 5.  
 GN CCR5.  
 OS Colobus guereza (Black-and-white colobus monkey).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Cercopithecidae; Colobinae;  
 OC Colobus.  
 OX NCBI\_TaxID=33548;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA Spearman P.W., Mburu D.N., Graham B.S.;  
 RT "Differential Utilization of CCR5 Molecules from Three East African  
 RT Simian Species by the HIV-1 Envelope Glycoprotein."  
 RL Submitted (APR-1999) to the EMBL/GenBank/DDBJ databases.  
 DR EMBL; AF141639; AAD32684.1; -.  
 DR GO; GO:0016021; C:integral to membrane; IEA.  
 DR GO; GO:0004872; F:receptor activity; IEA.  
 DR GO; GO:0001584; F:rhodopsin-like receptor activity; IEA.  
 DR GO; GO:0007186; P:G-protein coupled receptor protein signalin. . .; IEA.  
 DR InterPro; IPR000276; GPCR\_Rhodpsn.  
 DR Pfam; PF00001; 7tm\_1; 1.  
 DR PRINTS; PR00237; GPCRRHODOPSN.  
 DR PROSITE; PS00237; G\_PROTEIN\_RECEP\_F1\_1; 1.  
 DR PROSITE; PS50262; G\_PROTEIN\_RECEP\_F1\_2; 1.  
 KW Receptor.  
 SQ SEQUENCE 352 AA; 40550 MW; 9B078EF04D34DB36 CRC64;

Query Match 45.8%; Score 365; DB 6; Length 352;  
 Best Local Similarity 33.6%; Pred. No. 1.5e-32;  
 Matches 95; Conservative 10; Mismatches 28; Indels 150; Gaps 7;

Qy 1 MDYQVSSPIYDINYYTSEPCQKINVKQIAA----- 30  
 ||||| |||:|||||:|||||  
 Db 1 MDYQVSSPTYDIDYYTSEPCQKVNKQIAARLLPPLYSLVFIFGFVGNILVVLILINCKR 60  
 Qy 31 -----YKCGLC-----AAAQWDFGNTMCQHQRVHGHHSYKCG--- 64  
 | | ||||| | : | : |  
 Db 61 LKSMTDIYLLNLAISDLFFLLTVPFWAHYAAAQWDFGNTMC--QLLTGLYFIGFFSGIFF 118  
 Qy 65 -----LCTRSQKEGLHYTC 78  
 : |||:|||||  
 Db 119 IILLTIDRYLAIVHAVFALKARTATFGVVTSVITWVAVFASLPGIIFTRSQREGLHYTC 178  
 Qy 79 SSHFPYSQYQFWKNFQTLKI-----HQRVHGG----- 105  
 ||||| ||||| : : |  
 Db 179 SSHFPYSQYQFWKNFQTLKIVILGLVLP LLVMVICYS GILKTL LRCRNEKKRHRVRLIF 238  
 Qy 106 -----GGSYKCGLC----QEFGFLNNC SSSNRLDGHQRV 135  
 | | ||||| : |  
 Db 239 TIMIVYFLFWAPYNIVLLLNTFQEFGFLNNC SSSNRLDQAMQV 281

RESULT 14

Q95NC6

ID Q95NC6 PRELIMINARY; PRT; 352 AA.  
 AC Q95NC6;  
 DT 01-DEC-2001 (TrEMBLrel. 19, Created)  
 DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)  
 DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)  
 DE C-C chemokine receptor 5.  
 GN CCR5.  
 OS Trachypithecus johnii (hooded leaf monkey).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Cercopithecidae; Colobinae;  
 OC Trachypithecus.  
 OX NCBI\_TaxID=66063;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA Zhang Y., Ryder O.A., Zhang Y.;  
 RT "Sequence comparison of the CCR5 gene in primates and primate  
 RT phylogeny.";  
 RL Submitted (AUG-1999) to the EMBL/GenBank/DDBJ databases.  
 DR EMBL; AF177883; AAK43366.1; -.  
 DR GO; GO:0016021; C:integral to membrane; IEA.  
 DR GO; GO:0004872; F:receptor activity; IEA.  
 DR GO; GO:0001584; F:rhodopsin-like receptor activity; IEA.  
 DR GO; GO:0007186; P:G-protein coupled receptor protein signalin. . .; IEA.  
 DR InterPro; IPR000276; GPCR\_Rhodpsn.  
 DR Pfam; PF00001; 7tm\_1; 1.  
 DR PRINTS; PR00237; GPCRRHODOPSN.  
 DR PROSITE; PS00237; G\_PROTEIN\_RECEP\_F1\_1; 1.  
 DR PROSITE; PS50262; G\_PROTEIN\_RECEP\_F1\_2; 1.  
 KW Receptor.  
 SQ SEQUENCE 352 AA; 40462 MW; 52824E0322559F7F CRC64;

Query Match 45.8%; Score 365; DB 6; Length 352;

Best Local Similarity 33.6%; Pred. No. 1.5e-32;

Matches 95; Conservative 10; Mismatches 28; Indels 150; Gaps 7;

Qy 1 MDYQVSSPIYDINYTTSEPCQKINVKQIAA----- 30  
 ||||| |||:|||||||:|||||  
 Db 1 MDYQVSSPTYDIDYTTSEPCQKVNKQIAARLLPPLYSLVFIFGFVGNILVVLILINCKR 60  
 Qy 31 -----YKCGLC-----AAAQWDFGNTMCQHQRVHGHHSYKCG--- 64  
 | | ||||| | : | : |  
 Db 61 LKSMTDIYLLNLAISDLFFLLTVPFWAHYAAAQWDFGNTMC--QLLTGLYFIGFFSGIFF 118  
 Qy 65 -----LCTRSQKEGLHYTC 78  
 : |||:|||||  
 Db 119 IILLTIDRYLAIVHAVFALKARTVTFGVVTSVITWVAVFASLPGLIFTRSQREGLHYTC 178  
 Qy 79 SSHFPYSQYQFWKNFQTLKI-----HQRVHGG----- 105  
 ||||| ||||| : : |  
 Db 179 SSHFPYSQYQFWKNFQTLKIVILGLVLP LLVMVICYSGILKTLRLCRSEKKRHRVRLIF 238  
 Qy 106 -----GGSYKCGLC-----QEFFGLNNC SSSNRLDGHQRV 135  
 | | ||||| : |



## RESULT 15

Q95NC8

ID Q95NC8 PRELIMINARY; PRT; 352 AA.  
 AC Q95NC8;  
 DT 01-DEC-2001 (TrEMBLrel. 19, Created)  
 DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)  
 DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)  
 DE C-C chemokine receptor 5.  
 GN CCR5.  
 OS Colobus polykomos.  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Cercopithecidae; Colobinae;  
 OC Colobus.  
 OX NCBI\_TaxID=9572;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA Zhang Y., Ryder O.A., Zhang Y.;  
 RT "Sequence comparison of the CCR5 gene in primates and primate  
 RT phylogeny.";  
 RL Submitted (AUG-1999) to the EMBL/GenBank/DDBJ databases.  
 DR EMBL; AF177881; AAK43364.1; -.  
 DR GO; GO:0016021; C:integral to membrane; IEA.  
 DR GO; GO:0004872; F:receptor activity; IEA.  
 DR GO; GO:0001584; F:rhodopsin-like receptor activity; IEA.  
 DR GO; GO:0007186; P:G-protein coupled receptor protein signalin. . .; IEA.  
 DR InterPro; IPR000276; GPCR\_Rhodpsn.  
 DR Pfam; PF00001; 7tm\_1; 1.  
 DR PRINTS; PR00237; GPCRRHODOPSN.  
 DR PROSITE; PS00237; G\_PROTEIN\_RECEP\_F1\_1; 1.  
 DR PROSITE; PS50262; G\_PROTEIN\_RECEP\_F1\_2; 1.  
 KW Receptor.  
 SQ SEQUENCE 352 AA; 40578 MW; 4366F149C3B4938F CRC64;

Query Match 45.8%; Score 365; DB 6; Length 352;  
 Best Local Similarity 33.6%; Pred. No. 1.5e-32;  
 Matches 95; Conservative 10; Mismatches 28; Indels 150; Gaps 7;

Qy 1 MDYQVSSPIYDINYYTSEPCQKINVKQIAA----- 30  
 ||||| |||:|||||||:|||||  
 Db 1 MDYQVSSPTYDIDYYTSEPCQKVNKQIAARLLPPLYSLVFIFGFVGNILVVLILINCKR 60  
 Qy 31 -----YKCGLC-----AAAQWDFGNTMCQHQRVHGHHHSYKCG--- 64  
 | | ||||| | : | : |  
 Db 61 LKSMTDIYLLNLAISDLFFLLTVPFWAHYAAAQWDFGNTMC--QLLTGLYFIGFFSGIFF 118  
 Qy 65 -----LCTRSQKEGLHYTC 78  
 : ||||:|||||  
 Db 119 IILLTIDRYLAIVHAVFALKARTVTFGVVTSVITWVAVFASLPGIIFTRSQREGLHYTC 178  
 Qy 79 SSHFPYSQYQFWKNFQTLKI-----HQRVHGG----- 105  
 ||||| ||||| : : |  
 Db 179 SSHFPYSQYQFWKNFQTLKIVILGLVPLLVVICYSGLKTLRLCRNEKKRHRVRLIF 238  
 Qy 106 -----GGSYKCGLC-----QEFFGLNNCSSSNRLDGHQRV 135

Search completed: September 17, 2004, 16:48:12  
Job time : 68.8949 secs

OM protein - protein search, using sw model

Run on: September 17, 2004, 16:42:06 ; Search time 14.0339 Seconds  
(without alignments)  
512.023 Million cell updates/sec

Title: US-10-057-890A-10  
Perfect score: 797  
Sequence: 1 MDYQVSSPIYDINYYTSEPC.....GLNNCSSSNRLDGHQRVHAA 138

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 141681 seqs, 52070155 residues

Total number of hits satisfying chosen parameters: 141681

Minimum DB seq length: 0  
Maximum DB seq length: 20000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : SwissProt\_42:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	% Match	Query Length	DB	ID	Description
1	379	47.6	352	1	CKR5_HUMAN	P51681 homo sapien
2	374	46.9	352	1	CKR5_PANTR	P56440 pan troglod
3	369	46.3	352	1	CKR5_HYLSY	Q95nc5 hylobates s
4	369	46.3	352	1	CKR5_PONPY	O97881 pongo pygma
5	366	45.9	352	1	CKR5_PAPHA	P56441 papio hamad
6	365	45.8	352	1	CKR5_PYGBI	O97880 pygathrix b
7	365	45.8	352	1	CKR5_PYGNE	O97882 pygathrix n
8	365	45.8	352	1	CKR5_TRAFR	O97878 trachypithe
9	365	45.8	352	1	CKR5_TRAPH	O97879 trachypithe
10	364	45.7	352	1	CKR5_GORGO	P56439 gorilla gor
11	363	45.5	352	1	CKR5_MACMU	P79436 macaca mula
12	362	45.4	352	1	CKR5_HYMLL	Q95nc0 hylobates m
13	361	45.3	352	1	CKR5_CERPYP	Q9tv42 cercopithec
14	361	45.3	352	1	CKR5_CERTO	O62743 cercocebus
15	359	45.0	352	1	CKR5_HYLLE	O97883 hylobates l
16	357	44.8	352	1	CKR5_CERAE	P56493 cercopithec
17	272	34.1	354	1	CKR5_MOUSE	P51682 mus musculu

18	249	31.2	354	1	CKR5_RAT	O08556	rattus norv
19	156.5	19.6	474	1	Z256_HUMAN	Q9y2p7	homo sapien
20	154	19.3	201	1	Z239_MOUSE	P24399	mus musculu
21	151.5	19.0	645	1	Z235_MOUSE	Q61116	mus musculu
22	151.5	19.0	754	1	Z287_HUMAN	Q9hbt7	homo sapien
23	150.5	18.9	501	1	ZF96_MOUSE	Q9zld7	mus musculu
24	146.5	18.4	458	1	Z239_HUMAN	Q16600	homo sapien
25	144.5	18.1	759	1	Z287_MOUSE	Q9eqb9	mus musculu
26	144	18.1	698	1	Z234_HUMAN	Q14588	homo sapien
27	140.5	17.6	604	1	Z305_HUMAN	O43309	homo sapien
28	140.5	17.6	744	1	YJ62_HUMAN	Q8tf39	homo sapien
29	140.5	17.6	803	1	Z226_HUMAN	Q9nyt6	homo sapien
30	140.5	17.6	913	1	Z228_HUMAN	Q9uju3	homo sapien
31	139.5	17.5	578	1	Z192_HUMAN	Q15776	homo sapien
32	139.5	17.5	682	1	ZN45_HUMAN	Q02386	homo sapien
33	139.5	17.5	751	1	Z184_HUMAN	Q99676	homo sapien
34	139	17.4	810	1	Z33A_HUMAN	Q06730	homo sapien
35	138.5	17.4	468	1	ZF90_HUMAN	Q8tf47	homo sapien
36	138.5	17.4	670	1	ZN16_HUMAN	P17020	homo sapien
37	138.5	17.4	1029	1	Z197_HUMAN	O14709	homo sapien
38	138	17.3	463	1	Z331_HUMAN	Q9nqx6	homo sapien
39	136.5	17.1	348	1	Z134_HUMAN	P52741	homo sapien
40	136.5	17.1	659	1	Z304_HUMAN	Q9hcx3	homo sapien
41	136.5	17.1	825	1	ZF28_MOUSE	P10078	mus musculu
42	135.5	17.0	643	1	ZN74_HUMAN	Q16587	homo sapien
43	133.5	16.8	636	1	ZF90_MOUSE	Q61967	mus musculu
44	132.5	16.6	642	1	ZN14_HUMAN	P17017	homo sapien
45	131.5	16.5	367	1	Z211_HUMAN	Q13398	homo sapien

# ALIGNMENTS

## RESULT 1

### CKR5\_HUMAN

ID CKR5\_HUMAN STANDARD; PRT; 352 AA.

AC P51681; O14692; O14693; O14695; O14696; O14697; O14698; O14699;

AC O14700; O14701; O14702; O14703; O14704; O14705; O14706; O14707;

AC O14708; O15538; Q9UPA4;

DT 01-OCT-1996 (Rel. 34, Created)

DT 01-OCT-1996 (Rel. 34, Last sequence update)

DT 15-MAR-2004 (Rel. 43, Last annotation update)

DE C-C chemokine receptor type 5 (C-C CKR-5) (CC-CKR-5) (CCR-5) (CCR5)

DE (HIV-1 fusion co-receptor) (CHEMR13) (CD195 antigen).

GN CCR5 OR CMKBR5.

OS Homo sapiens (Human).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

OX NCBI\_TaxID=9606;

RN [1]

RP SEQUENCE FROM N.A.

RX MEDLINE=96241590; PubMed=8639485;

RA Samson M., Labbe O., Mollereau C., Vassart G., Parmentier M.;

RT "Molecular cloning and functional expression of a new human

RT CC-chemokine receptor gene.";

RL Biochemistry 35:3362-3367(1996).

RN [2]

RP SEQUENCE FROM N.A.  
 RX MEDLINE=96291862; PubMed=8663314;  
 RA Raport C.J., Gosling J., Schweichart V.L., Gray P.W., Charo I.F.;  
 RT "Molecular cloning and functional characterization of a novel human  
 RT CC chemokine receptor (CCR5) for RANTES, MIP-1beta, and MIP-1alpha.";  
 RL J. Biol. Chem. 271:17161-17166(1996).  
 RN [3]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=96295970; PubMed=8699119;  
 RA Combadiere C., Ahuja S.K., Tiffany H.L., Murphy P.M.;  
 RT "Cloning and functional expression of CC CKR5, a human monocyte CC  
 RT chemokine receptor selective for MIP-1(alpha), MIP-1(beta), and  
 RT RANTES.";  
 RL J. Leukoc. Biol. 60:147-152(1996).  
 RN [4]  
 RP SEQUENCE FROM N.A.  
 RA McCombie W.R., Wilson R., Chen E., Gibbs R., Zuo L., Johnson D.,  
 RA Nhan M., Parnell L., Dedhia N., Ansari A., Mardis E., Schutz K.,  
 RA Gnoj L., la Bastide M., Kaplan N., Greco T., Touchman J.,  
 RA Muzny D., Chen C.N., Evans C., Fitzgerald M., See L.H., Tang M.,  
 RA Porcel B.M., Dragan Y., Giacalone J., Pae A., Powell E.,  
 RA Solinsky K.A., Desilva U., Diaz-Perez S., Zhou X., Yu Y.,  
 RA Watanabe M., Doggett N., Garcia D., Sagripanti J.L.;  
 RL Submitted (MAY-1997) to the EMBL/GenBank/DDBJ databases.  
 RN [5]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=98001387; PubMed=9343222;  
 RA Kuhmann S.E., Platt E.J., Kozak S.L., Kabat D.;  
 RT "Polymorphisms in the CCR5 genes of African green monkeys and mice  
 RT implicate specific amino acids in infections by simian and human  
 RT immunodeficiency viruses.";  
 RL J. Virol. 71:8642-8656(1997).  
 RN [6]  
 RP SEQUENCE FROM N.A., AND POLYMORPHISMS.  
 RX MEDLINE=98022612; PubMed=9359654;  
 RA Zhang L., Carruthers C.D., He T., Huang Y., Cao Y., Wang G., Hahn B.,  
 RA Ho D.D.;  
 RT "HIV type 1 subtypes, coreceptor usage, and CCR5 polymorphism.";  
 RL AIDS Res. Hum. Retroviruses 13:1357-1366(1997).  
 RN [7]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=98049523; PubMed=9388201;  
 RA Mummidi S., Ahuja S.S., McDaniel B.L., Ahuja S.K.;  
 RT "The human CC chemokine receptor 5 (CCR5) gene. Multiple transcripts  
 RT with 5'-end heterogeneity, dual promoter usage, and evidence for  
 RT polymorphisms within the regulatory regions and noncoding exons.";  
 RL J. Biol. Chem. 272:30662-30671(1997).  
 RN [8]  
 RP SEQUENCE FROM N.A., AND VARIANT ARG-178.  
 RA Magierowska M., Barre-Sinoussi F., Issafras H., Theodorou I.,  
 RA Debre P.;  
 RL Submitted (MAR-1998) to the EMBL/GenBank/DDBJ databases.  
 RN [9]  
 RP SEQUENCE FROM N.A.  
 RA Kopatz S.A., Aronstam R.S., Sharma S.V.;  
 RT "cDNA clones of human proteins involved in signal transduction  
 RT sequenced by the Guthrie cDNA resource center (www.cdna.org).";

RL Submitted (JAN-2003) to the EMBL/GenBank/DDBJ databases.  
 RN [10]  
 RP CHARACTERIZATION OF ITS HIV-1 RECEPTOR FUNCTION.  
 RX MEDLINE=96260017; PubMed=8649511;  
 RA Deng H., Liu R., Ellmeier W., Choe S., Unutmaz D., Burkhart M.,  
 RA di Marzio P., Marmon S., Sutton R.E., Hill C.M., Davis C.B.,  
 RA Peiper S.C., Schall T.J., Littman D.R., Landau N.R.;  
 RT "Identification of a major co-receptor for primary isolates of  
 RT HIV-1.";  
 RL Nature 381:661-666(1996).  
 RN [11]  
 RP CHARACTERIZATION OF ITS HIV-1 RECEPTOR FUNCTION.  
 RX MEDLINE=96260018; PubMed=8649512;  
 RA Dragic T., Litwin V., Allaway G.P., Martin S.R., Huang Y.,  
 RA Nagashima K.A., Cayanan C., Maddon P.J., Koup R.A., Moore J.P.,  
 RA Paxton W.A.;  
 RT "HIV-1 entry into CD4+ cells is mediated by the chemokine receptor  
 RT CC-CKR-5.";  
 RL Nature 381:667-673(1996).  
 RN [12]  
 RP SULFATION.  
 RX MEDLINE=99189752; PubMed=10089882;  
 RA Farzan M., Mirzabekov T., Kolchinsky P., Wyatt R., Cayabyab M.,  
 RA Gerard N.P., Gerard C., Sodroski J., Choe H.;  
 RT "Tyrosine sulfation of the amino terminus of CCR5 facilitates HIV-1  
 RT entry.";  
 RL Cell 96:667-676(1999).  
 CC -!- FUNCTION: Receptor for a C-C type chemokine. Binds to MIP-1-alpha,  
 CC MIP-1-beta and rantes and subsequently transduces a signal by  
 CC increasing the intracellular calcium ions level. May play a role  
 CC in the control of granulocytic lineage proliferation or  
 CC differentiation. Acts as co-receptor with CD4 for primary non-  
 CC syncytium-inducing strains (NSI) (macrophage-tropic) of HIV-1  
 CC virus. It promotes Env-mediated fusion of the virus.  
 CC -!- SUBCELLULAR LOCATION: Integral membrane protein.  
 CC -!- TISSUE SPECIFICITY: Found in promyelocytic cells.  
 CC -!- PTM: Sulfation contributes to the efficiency of HIV-1 entry.  
 CC -!- PTM: Modified by O-linked glycosylation, but not by N-linked  
 CC glycosylation.  
 CC -!- SIMILARITY: Belongs to family 1 of G-protein coupled receptors.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 DR EMBL; X91492; CAA62796.1; -.  
 DR EMBL; U54994; AAC50598.1; -.  
 DR EMBL; U57840; AAB17071.1; -.  
 DR EMBL; U95626; AAB57793.1; -.  
 DR EMBL; U83326; AAC51797.1; -.  
 DR EMBL; AF011500; AAB65700.1; -.  
 DR EMBL; AF011501; AAB65701.1; -.  
 DR EMBL; AF011502; AAB65702.1; -.

DR EMBL; AF011503; AAB65703.1; -.  
 DR EMBL; AF011505; AAB65705.1; -.  
 DR EMBL; AF011506; AAB65706.1; -.  
 DR EMBL; AF011507; AAB65707.1; -.  
 DR EMBL; AF011508; AAB65708.1; -.  
 DR EMBL; AF011509; AAB65709.1; -.  
 DR EMBL; AF011510; AAB65710.1; -.  
 DR EMBL; AF011511; AAB65711.1; -.  
 DR EMBL; AF011512; AAB65712.1; -.  
 DR EMBL; AF011513; AAB65713.1; -.  
 DR EMBL; AF011514; AAB65714.1; -.  
 DR EMBL; AF011515; AAB65715.1; -.  
 DR EMBL; AF011516; AAB65716.1; -.  
 DR EMBL; AF011517; AAB65717.1; -.  
 DR EMBL; AF011518; AAB65718.1; -.  
 DR EMBL; AF011519; AAB65719.1; -.  
 DR EMBL; AF011520; AAB65720.1; -.  
 DR EMBL; AF011521; AAB65721.1; -.  
 DR EMBL; AF011522; AAB65722.1; -.  
 DR EMBL; AF011523; AAB65723.1; -.  
 DR EMBL; AF011524; AAB65724.1; -.  
 DR EMBL; AF011525; AAB65725.1; -.  
 DR EMBL; AF011526; AAB65726.1; -.  
 DR EMBL; AF011527; AAB65727.1; -.  
 DR EMBL; AF011528; AAB65728.1; -.  
 DR EMBL; AF011529; AAB65729.1; -.  
 DR EMBL; AF011530; AAB65730.1; -.  
 DR EMBL; AF011531; AAB65731.1; -.  
 DR EMBL; AF011532; AAB65732.1; -.  
 DR EMBL; AF011533; AAB65733.1; -.  
 DR EMBL; AF011534; AAB65734.1; -.  
 DR EMBL; AF011535; AAB65735.1; -.  
 DR EMBL; AF011536; AAB65736.1; -.  
 DR EMBL; AF011537; AAB65737.1; -.  
 DR EMBL; AF031237; AAB94735.1; -.  
 DR EMBL; AF052539; AAD18131.1; -.  
 DR EMBL; AY221093; AAO65971.1; -.  
 DR Genew; HGNC:1606; CCR5.  
 DR MIM; 601373; -.  
 DR GO; GO:0005768; C:endosome; TAS.  
 DR GO; GO:0005887; C:integral to plasma membrane; TAS.  
 DR GO; GO:0016493; F:C-C chemokine receptor activity; NAS.  
 DR GO; GO:0015026; F:coreceptor activity; TAS.  
 DR GO; GO:0007267; P:cell-cell signaling; TAS.  
 DR GO; GO:0006968; P:cellular defense response; TAS.  
 DR GO; GO:0006935; P:chemotaxis; TAS.  
 DR GO; GO:0007204; P:cytosolic calcium ion concentration elevation; TAS.  
 DR GO; GO:0007186; P:G-protein coupled receptor protein signalin. . .; TAS.  
 DR GO; GO:0006954; P:inflammatory response; TAS.  
 DR GO; GO:0007125; P:invasive growth; TAS.  
 DR GO; GO:0007203; P:phosphatidylinositol-4,5-bisphosphate hydro. . .; TAS.  
 DR InterPro; IPR000276; GPCR\_Rhodpsn.  
 DR Pfam; PF00001; 7tm\_1; 1.  
 DR PRINTS; PR00237; GPCRRHODOPSN.  
 DR PROSITE; PS00237; G\_PROTEIN\_RECEP\_F1\_1; 1.  
 DR PROSITE; PS50262; G\_PROTEIN\_RECEP\_F1\_2; 1.  
 KW G-protein coupled receptor; Transmembrane; Glycoprotein; Sulfation;

KW	Polymorphism.			
FT	DOMAIN	1	30	EXTRACELLULAR (POTENTIAL).
FT	TRANSMEM	31	58	1 (POTENTIAL).
FT	DOMAIN	59	68	CYTOPLASMIC (POTENTIAL).
FT	TRANSMEM	69	89	2 (POTENTIAL).
FT	DOMAIN	90	102	EXTRACELLULAR (POTENTIAL).
FT	TRANSMEM	103	124	3 (POTENTIAL).

Query Match 47.6%; Score 379; DB 1; Length 352;  
 Best Local Similarity 35.0%; Pred. No. 1.8e-29;  
 Matches 99; Conservative 7; Mismatches 27; Indels 150; Gaps 7;

Qy	1	MDYQVSSPIYDINYYTSEPCQKINVKQIAA-----	30
Db	1	MDYQVSSPIYDINYYTSEPCQKINVKQIAARLLPPLYSLVFIFGFVGNMLVILILINCKR	60
Qy	31	-----YKCGLC-----AAAQWDFGNTMCQHQRVHGHHSYKCG---	64
		:   :	
Db	61	LKSMTDIYLLNLAISDLFFLLTVPFWAHYAAAQWDFGNTMC--QLLTGLYFIGFFSGIFF	118
Qy	65	-----LCTRSQKEGLHYTC	78
		:	
Db	119	IILLTIDRYLAVVHAVFALKARTVTFGVVTSVITWVVAVFASLPGIIFTRSQKEGLHYTC	178
Qy	79	SSHFPYSQYQFWKNFQTLKI-----HQRVHGG-----	105
		: :	
Db	179	SSHFPYSQYQFWKNFQTLKIVILGLVLPPLVMVICYSGLKTLRLCRNEKKRHRVRLIF	238
Qy	106	-----GGSYKCGLC---QEFGNLCSSSNRLDGHQRV	135
		:	
Db	239	TIMIVYFLFWAPYNIVLLLNTFQEFGNLCSSSNRLDQAMQV	281

## RESULT 2

### CKR5\_PANTR

ID CKR5\_PANTR STANDARD; PRT; 352 AA.  
 AC P56440; O02778;  
 DT 15-JUL-1998 (Rel. 36, Created)  
 DT 15-JUL-1998 (Rel. 36, Last sequence update)  
 DT 16-OCT-2001 (Rel. 40, Last annotation update)  
 DE C-C chemokine receptor type 5 (C-C CKR-5) (CC-CKR-5) (CCR-5) (CCR5).  
 GN CCR5 OR CMKBR5.  
 OS Pan troglodytes (Chimpanzee).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Pan.  
 OX NCBI\_TaxID=9598;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=97268687; PubMed=9108095;  
 RA Edinger A.L., Amedee A., Miller K., Doranz B.J., Endres M.,  
 RA Sharron M., Samson M., Lu Z.H., Clements J.E., Murphey-Corb M.,  
 RA Peiper S.C., Parmentier M., Broder C.C., Doms R.W.;  
 RT "Differential utilization of CCR5 by macrophage and T cell tropic  
 RT simian immunodeficiency virus strains."  
 RL Proc. Natl. Acad. Sci. U.S.A. 94:4005-4010(1997).  
 RN [2]  
 RP SEQUENCE FROM N.A.



RA Zimmerman P.A., Buckler-White A., Alkhatib G.;  
 RL Submitted (MAR-1997) to the EMBL/GenBank/DDBJ databases.  
 RN [3]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=98022612; PubMed=9359654;  
 RA Zhang L., Carruthers C.D., He T., Huang Y., Cao Y., Wang G., Hahn B.,  
 RA Ho D.D.;  
 RT "HIV type 1 subtypes, coreceptor usage, and CCR5 polymorphism";  
 RL AIDS Res. Hum. Retroviruses 13:1357-1366(1997).  
 RN [4]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=97426118; PubMed=9282822;  
 RA Zacharova V., Zachar V., Goustin A.S.;  
 RT "Sequence of chemokine receptor gene CCR5 in chimpanzees, a natural  
 RT HIV type 1 host.";  
 RL AIDS Res. Hum. Retroviruses 13:1159-1161(1997).  
 RN [5]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=98090115; PubMed=9430250;  
 RA Pretet J.-L., Zerbib A., Girard M., Guillet J.-G., Butor C.;  
 RT "Chimpanzee CXCR4 and CCR5 act as coreceptors for HIV type 1.";  
 RL AIDS Res. Hum. Retroviruses 13:1583-1587(1997).  
 RN [6]  
 RP SEQUENCE FROM N.A.  
 RA Zhang Y., Ryder O.A., Zhang Y.;  
 RT "Sequence comparison of the CCR5 gene in primates and primate  
 RT phylogeny.";  
 RL Submitted (AUG-1999) to the EMBL/GenBank/DDBJ databases.  
 CC -!- FUNCTION: Receptor for a C-C type chemokine. Binds to MIP-1-alpha,  
 CC MIP-1-beta and RANTES and subsequently transduces a signal by  
 CC increasing the intracellular calcium ions level. May play a role  
 CC in the control of granulocytic lineage proliferation or  
 CC differentiation.  
 CC -!- SUBCELLULAR LOCATION: Integral membrane protein.  
 CC -!- SIMILARITY: Belongs to family 1 of G-protein coupled receptors.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 DR EMBL; AF005663; AAB62557.1; -.  
 DR EMBL; U94329; AAB58446.1; -.  
 DR EMBL; AF011542; AAB65742.1; -.  
 DR EMBL; U97666; AAC51670.1; -.  
 DR EMBL; AF011540; AAB65740.1; -.  
 DR EMBL; U89797; AAC03717.1; -.  
 DR EMBL; AF177894; AAK43377.1; -.  
 DR InterPro; IPR000276; GPCR\_Rhodpsn.  
 DR Pfam; PF00001; 7tm\_1; 1.  
 DR PRINTS; PR00237; GPCRRHODOPSN.  
 DR PROSITE; PS00237; G\_PROTEIN\_RECEP\_F1\_1; 1.  
 DR PROSITE; PS50262; G\_PROTEIN\_RECEP\_F1\_2; 1.  
 KW G-protein coupled receptor; Transmembrane; Glycoprotein; Sulfation.

DE C-C chemokine receptor type 5 (C-C CKR-5) (CC-CKR-5) (CCR-5) (CCR5).

GN CCR5 OR CMKBR5.  
 OS Hylobates syndactylus (Siamang) (Symphalangus syndactylus).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Hylobatidae; Hylobates.  
 OX NCBI\_TaxID=9590;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=99416438; PubMed=10486970;  
 RA Zhang Y.-W., Ryder O.A., Zhang Y.-P.;  
 RT "Sequence evolution of the CCR5 chemokine receptor gene in primates.";  
 RL Mol. Biol. Evol. 16:1145-1154(1999).  
 CC -!- FUNCTION: Receptor for a C-C type chemokine. Binds to MIP-1-alpha,  
 CC MIP-1-beta and RANTES and subsequently transduces a signal by  
 CC increasing the intracellular calcium ions level. May play a role  
 CC in the control of granulocytic lineage proliferation or  
 CC differentiation.  
 CC -!- SUBCELLULAR LOCATION: Integral membrane protein.  
 CC -!- SIMILARITY: Belongs to family 1 of G-protein coupled receptors.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 DR EMBL; AF177884; AAK43367.1; -.  
 DR InterPro; IPR000276; GPCR\_Rhodpsn.  
 DR Pfam; PF00001; 7tm\_1; 1.  
 DR PRINTS; PR00237; GPCRRHODOPSN.  
 DR PROSITE; PS00237; G\_PROTEIN\_RECEP\_F1\_1; 1.  
 DR PROSITE; PS50262; G\_PROTEIN\_RECEP\_F1\_2; 1.  
 KW G-protein coupled receptor; Transmembrane; Glycoprotein; Sulfation.  
 FT DOMAIN 1 30 EXTRACELLULAR (POTENTIAL).  
 FT TRANSMEM 31 58 1 (POTENTIAL).  
 FT DOMAIN 59 68 CYTOPLASMIC (POTENTIAL).  
 FT TRANSMEM 69 89 2 (POTENTIAL).  
 FT DOMAIN 90 102 EXTRACELLULAR (POTENTIAL).  
 FT TRANSMEM 103 124 3 (POTENTIAL).  
 FT DOMAIN 125 141 CYTOPLASMIC (POTENTIAL).  
 FT TRANSMEM 142 166 4 (POTENTIAL).  
 FT DOMAIN 167 198 EXTRACELLULAR (POTENTIAL).  
 FT TRANSMEM 199 218 5 (POTENTIAL).  
 FT DOMAIN 219 235 CYTOPLASMIC (POTENTIAL).  
 FT TRANSMEM 236 260 6 (POTENTIAL).  
 FT DOMAIN 261 277 EXTRACELLULAR (POTENTIAL).  
 FT TRANSMEM 278 301 7 (POTENTIAL).  
 FT DOMAIN 302 352 CYTOPLASMIC (POTENTIAL).  
 FT DISULFID 101 178 BY SIMILARITY.  
 FT MOD\_RES 3 3 SULFATION (BY SIMILARITY).  
 FT MOD\_RES 10 10 SULFATION (BY SIMILARITY).  
 FT MOD\_RES 14 14 SULFATION (BY SIMILARITY).  
 SQ SEQUENCE 352 AA; 40508 MW; F4F64B3AD5AF658A CRC64;

Query Match 46.3%; Score 369; DB 1; Length 352;  
 Best Local Similarity 34.3%; Pred. No. 1.7e-28;

Matches 97; Conservative 8; Mismatches 28; Indels 150; Gaps 7;

```

Qy      1 MDYQVSSPIYDINYYTSEPCQKINVKQIAA----- 30
      ||||| |||:|||||
Db      1 MDYQVSSPTYDIDYYTSEPCQKINVKQIAARLLPPLYSLVFIFGFVGNMLVILILINCKR 60

Qy     31 -----YKCGLC-----AAAQWDFGNTMCQHQRVHGHHSYKCG--- 64
      | | ||||| | : | : |
Db     61 LKSMTDIYLLNLAISDLFFLLTVPFWAHYAAAQWDFGNTMC--QLLTGLYFIGFFSGIFF 118

Qy     65 -----LCTRSQKEGLHYTC 78
      : |||||
Db    119 IILLTIDRYLAIVHAVFALKARTVTFGVVTSVITWVVAVFASLPGIIFTRSQKEGLHYTC 178

Qy     79 SSHFPYSQYQFWKNFQTLKI-----HQRVHGG----- 105
      ||||| ||||| : |
Db    179 SSHFPYSQYQFWKNFQTLKIVILGLVLP LLVMVICYS GILKTL LRCRNEKKRHRAVRLIF 238

Qy    106 -----GGSYKCGLC----QEFG LNNC SSSNRLDGHQRV 135
      | | ||||| : |
Db    239 TIMIVYFLFWAPYNIVLLLNTFQEFG LNNC SSSNRLDQAMQV 281

```

#### RESULT 4

CKR5\_PONPY

ID CKR5\_PONPY STANDARD; PRT; 352 AA.

AC O97881;

DT 30-MAY-2000 (Rel. 39, Created)

DT 30-MAY-2000 (Rel. 39, Last sequence update)

DT 28-FEB-2003 (Rel. 41, Last annotation update)

DE C-C chemokine receptor type 5 (C-C CKR-5) (CC-CKR-5) (CCR-5) (CCR5).

GN CCR5 OR CMKBR5.

OS Pongo pygmaeus (Orangutan).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Pongo.

OX NCBI\_TaxID=9600;

RN [1]

RP SEQUENCE FROM N.A.

RX MEDLINE=99416438; PubMed=10486970;

RA Zhang Y.-W., Ryder O.A., Zhang Y.-P.;

RT "Sequence evolution of the CCR5 chemokine receptor gene in primates.";

RL Mol. Biol. Evol. 16:1145-1154(1999).

CC -!- FUNCTION: Receptor for a C-C type chemokine. Binds to MIP-1-alpha,  
 CC MIP-1-beta and RANTES and subsequently transduces a signal by  
 CC increasing the intracellular calcium ions level. May play a role  
 CC in the control of granulocytic lineage proliferation or  
 CC differentiation.

CC -!- SUBCELLULAR LOCATION: Integral membrane protein.

CC -!- SIMILARITY: Belongs to family 1 of G-protein coupled receptors.

CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

```

CC -----
DR EMBL; AF075446; AAD19858.1; -.
DR InterPro; IPR000276; GPCR_Rhodpsn.
DR Pfam; PF00001; 7tm_1; 1.
DR PRINTS; PR00237; GPCRRHODOPSN.
DR PROSITE; PS00237; G_PROTEIN_RECEP_F1_1; 1.
DR PROSITE; PS50262; G_PROTEIN_RECEP_F1_2; 1.
KW G-protein coupled receptor; Transmembrane; Glycoprotein; Sulfation.
FT DOMAIN 1 30 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 31 58 1 (POTENTIAL).
FT DOMAIN 59 68 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 69 89 2 (POTENTIAL).
FT DOMAIN 90 102 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 103 124 3 (POTENTIAL).
FT DOMAIN 125 141 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 142 166 4 (POTENTIAL).
FT DOMAIN 167 198 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 199 218 5 (POTENTIAL).
FT DOMAIN 219 235 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 236 260 6 (POTENTIAL).
FT DOMAIN 261 277 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 278 301 7 (POTENTIAL).
FT DOMAIN 302 352 CYTOPLASMIC (POTENTIAL).
FT DISULFID 101 178 BY SIMILARITY.
FT MOD_RES 3 3 SULFATION (BY SIMILARITY).
FT MOD_RES 10 10 SULFATION (BY SIMILARITY).
FT MOD_RES 14 14 SULFATION (BY SIMILARITY).
FT MOD_RES 15 15 SULFATION (BY SIMILARITY).
SQ SEQUENCE 352 AA; 40527 MW; F4E2F47135AF658A CRC64;

```

```

Query Match 46.3%; Score 369; DB 1; Length 352;
Best Local Similarity 34.3%; Pred. No. 1.7e-28;
Matches 97; Conservative 8; Mismatches 28; Indels 150; Gaps 7;

```

```

Qy 1 MDYQVSSPIYDINYYTSEPCQKINVQIAA----- 30
    ||||| |||:|||||
Db 1 MDYQVSSPTYDIDYYTSEPCQKINVQIAARLLPPLYSLVFIFGFVGNMLVILILINCKR 60

Qy 31 -----YKCGLC-----AAAQWDFGNTMCQHQRVHGHHSYKCG--- 64
    | | ||||| | : | : |
Db 61 LKSMTDIYLLNLAISDLFFLLTVPFWAHYAAAQWDFGNTMC--QLLTGLYFIGFFSGIFF 118

Qy 65 -----LCTRSQKEGLHYTC 78
    : |||||
Db 119 IILLTIDRYLAIVHAVFALKARTVTFGVVTSVITWVAVFASLPGIIFTRSQKEGLHYTC 178

Qy 79 SSHFPYSQYQFWKNFQTLKI-----HQRVHGG----- 105
    ||||| ||||| :|
Db 179 SSHFPYSQYQFWKNFQTLKIVILGLVLP LLVMVICYSGLKTL LRCRNEKKRHRAVRLIF 238

Qy 106 -----GGSYKCGLC----QEFGGLNNC SSSNRLDGHQRV 135
    | | ||||| :|
Db 239 TIMIVYFLWAPYNIVLLNTFQEFFGLNNC SSSNRLDQAMQV 281

```

RESULT 5  
 CKR5\_PAPHA

ID CKR5\_PAPHA STANDARD; PRT; 352 AA.  
 AC P56441;  
 DT 15-JUL-1998 (Rel. 36, Created)  
 DT 15-JUL-1998 (Rel. 36, Last sequence update)  
 DT 16-OCT-2001 (Rel. 40, Last annotation update)  
 DE C-C chemokine receptor type 5 (C-C CKR-5) (CC-CKR-5) (CCR-5) (CCR5).  
 GN CCR5 OR CMKBR5.  
 OS Papio hamadryas (Hamadryas baboon), and  
 OS Papio anubis (Olive baboon).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Cercopithecidae;  
 OC Cercopithecinae; Papio.  
 OX NCBI\_TaxID=9557, 9555;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC SPECIES=P.hamadryas;  
 RX MEDLINE=97268687; PubMed=9108095;  
 RA Edinger A.L., Amedee A., Miller K., Doranz B.J., Endres M.,  
 RA Sharron M., Samson M., Lu Z.H., Clements J.E., Murphey-Corb M.,  
 RA Peiper S.C., Parmentier M., Broder C.C., Doms R.W.;  
 RT "Differential utilization of CCR5 by macrophage and T cell tropic  
 RT simian immunodeficiency virus strains.";  
 RL Proc. Natl. Acad. Sci. U.S.A. 94:4005-4010(1997).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RC SPECIES=P.hamadryas;  
 RX MEDLINE=99210133; PubMed=10195758;  
 RA Saksena N.K., Wang B., Novembre F.J., Bolton W., Smit T.K., Lal R.B.;  
 RT "Species-specific changes in the CCR5 gene from African and Asian  
 RT nonhuman primates.";  
 RL AIDS Res. Hum. Retroviruses 15:479-483(1999).  
 RN [3]  
 RP SEQUENCE FROM N.A.  
 RC SPECIES=P.anubis;  
 RA Benton P.A., Timanus D.K., Shearer M.H., Lee D.R., Kennedy R.C.;  
 RL Submitted (NOV-1997) to the EMBL/GenBank/DDBJ databases.  
 CC -!- FUNCTION: Receptor for a C-C type chemokine. Binds to MIP-1-alpha,  
 CC MIP-1-beta and RANTES and subsequently transduces a signal by  
 CC increasing the intracellular calcium ions level. May play a role  
 CC in the control of granulocytic lineage proliferation or  
 CC differentiation.  
 CC -!- SUBCELLULAR LOCATION: Integral membrane protein.  
 CC -!- SIMILARITY: Belongs to family 1 of G-protein coupled receptors.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 DR EMBL; AF005658; AAB62552.1; -.  
 DR EMBL; AF105287; AAD20556.1; -.  
 DR EMBL; AF105288; AAD20557.1; -.  
 DR EMBL; AF105289; AAD20558.1; -.  
 DR EMBL; AF105290; AAD20559.1; -.

DR EMBL; AF023452; AAC63830.1; -.  
DR InterPro; IPR000276; GPCR\_Rhodpsn.  
DR Pfam; PF00001; 7tm\_1; 1.  
DR PRINTS; PR00237; GPCRRHODOPSN.  
DR PROSITE; PS00237; G\_PROTEIN\_RECEP\_F1\_1; 1.  
DR PROSITE; PS50262; G\_PROTEIN\_RECEP\_F1\_2; 1.  
KW G-protein coupled receptor; Transmembrane; Glycoprotein; Sulfation.  
FT DOMAIN 1 30 EXTRACELLULAR (POTENTIAL).  
FT TRANSMEM 31 58 1 (POTENTIAL).  
FT DOMAIN 59 68 CYTOPLASMIC (POTENTIAL).  
FT TRANSMEM 69 89 2 (POTENTIAL).  
FT DOMAIN 90 102 EXTRACELLULAR (POTENTIAL).  
FT TRANSMEM 103 124 3 (POTENTIAL).  
FT DOMAIN 125 141 CYTOPLASMIC (POTENTIAL).  
FT TRANSMEM 142 166 4 (POTENTIAL).  
FT DOMAIN 167 198 EXTRACELLULAR (POTENTIAL).  
FT TRANSMEM 199 218 5 (POTENTIAL).  
FT DOMAIN 219 235 CYTOPLASMIC (POTENTIAL).  
FT TRANSMEM 236 260 6 (POTENTIAL).  
FT DOMAIN 261 277 EXTRACELLULAR (POTENTIAL).  
FT TRANSMEM 278 301 7 (POTENTIAL).  
FT DOMAIN 302 352 CYTOPLASMIC (POTENTIAL).  
FT DISULFID 101 178 BY SIMILARITY.  
FT MOD\_RES 3 3 SULFATION (BY SIMILARITY).  
FT MOD\_RES 10 10 SULFATION (BY SIMILARITY).  
FT MOD\_RES 14 14 SULFATION (BY SIMILARITY).  
FT MOD\_RES 15 15 SULFATION (BY SIMILARITY).  
FT CARBOHYD 268 268 N-LINKED (GLCNAC. . .) (POTENTIAL).  
SQ SEQUENCE 352 AA; 40489 MW; 5E1504A9BA1FE8B2 CRC64;

Query Match 45.9%; Score 366; DB 1; Length 352;

Best Local Similarity 33.9%; Pred. No. 3.3e-28;

Matches 96; Conservative 9; Mismatches 28; Indels 150; Gaps 7;

```

Qy      1 MDYQVSSPIYDINYYTSEPCQKINVKQIAA----- 30
          ||||| |||:||||| ||||| |||||
Db      1 MDYQVSSPTYDIDYYTSEPCQKINVKQIAARLLPPLYSLVFIFGFVGNILVVLILINCKR 60

Qy     31 -----YKCGLC-----AAAQWDFGNTMCQHQRVHGHHHSYKCG--- 64
          |  |                      ||||| ||| | : | : |
Db     61 LKSMTDIYLLNLAISDLLFLLTVPFWAHYAAAQWDFGNTMC--QLLTGLYFIGFFSGIFF 118

Qy     65 -----LCTRSQKEGLHYTC 78
          : ||||:|||||
Db    119 IILLTIDRYLAIVHAVFALKARTVTFGVVTSVITWVVAVFASLPGLIIFTRSQREGLHYTC 178

Qy     79 SSHFPYSQYQFWKNFQTLKI-----HQRVHGG----- 105
          ||||| ||||| ||||| ||||| :|
Db    179 SSHFPYSQYQFWKNFQTLKIVILGLVLP LLVMVICYSGLKTL LRCRNEKKRHRVRLIF 238

Qy    106 -----GGSYKCGLC----QEFFGLNNC SSSNRLDGHQRV 135
          |  |                      ||||| ||||| ||||| :|
Db    239 TIMIVYFLFWAPYNIVLLNTFQEFFGLNNC SSSNRLDQAMQV 281

```

RESULT 6  
CKR5\_PYGBI

ID CKR5\_PYGBI STANDARD; PRT; 352 AA.  
 AC 097880;  
 DT 30-MAY-2000 (Rel. 39, Created)  
 DT 30-MAY-2000 (Rel. 39, Last sequence update)  
 DT 28-FEB-2003 (Rel. 41, Last annotation update)  
 DE C-C chemokine receptor type 5 (C-C CKR-5) (CC-CKR-5) (CCR-5) (CCR5).  
 GN CCR5 OR CMKBR5.  
 OS *Pygathrix bieti* (Black snub-nosed monkey) (*Rhinopithecus bieti*).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Cercopithecidae; Colobinae;  
 OC *Pygathrix*.  
 OX NCBI\_TaxID=61621;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=99416438; PubMed=10486970;  
 RA Zhang Y.-W., Ryder O.A., Zhang Y.-P.;  
 RT "Sequence evolution of the CCR5 chemokine receptor gene in primates.";  
 RL Mol. Biol. Evol. 16:1145-1154(1999).  
 CC -!- FUNCTION: Receptor for a C-C type chemokine. Binds to MIP-1-alpha,  
 CC MIP-1-beta and RANTES and subsequently transduces a signal by  
 CC increasing the intracellular calcium ions level. May play a role  
 CC in the control of granulocytic lineage proliferation or  
 CC differentiation.  
 CC -!- SUBCELLULAR LOCATION: Integral membrane protein.  
 CC -!- SIMILARITY: Belongs to family 1 of G-protein coupled receptors.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 DR EMBL; AF075445; AAD19857.1; -.  
 DR InterPro; IPR000276; GPCR\_Rhodpsn.  
 DR Pfam; PF00001; 7tm\_1; 1.  
 DR PRINTS; PR00237; GPCRRHODOPSN.  
 DR PROSITE; PS00237; G\_PROTEIN\_RECEP\_F1\_1; 1.  
 DR PROSITE; PS50262; G\_PROTEIN\_RECEP\_F1\_2; 1.  
 KW G-protein coupled receptor; Transmembrane; Glycoprotein; Sulfation.  
 FT DOMAIN 1 30 EXTRACELLULAR (POTENTIAL).  
 FT TRANSMEM 31 58 1 (POTENTIAL).  
 FT DOMAIN 59 68 CYTOPLASMIC (POTENTIAL).  
 FT TRANSMEM 69 89 2 (POTENTIAL).  
 FT DOMAIN 90 102 EXTRACELLULAR (POTENTIAL).  
 FT TRANSMEM 103 124 3 (POTENTIAL).  
 FT DOMAIN 125 141 CYTOPLASMIC (POTENTIAL).  
 FT TRANSMEM 142 166 4 (POTENTIAL).  
 FT DOMAIN 167 198 EXTRACELLULAR (POTENTIAL).  
 FT TRANSMEM 199 218 5 (POTENTIAL).  
 FT DOMAIN 219 235 CYTOPLASMIC (POTENTIAL).  
 FT TRANSMEM 236 260 6 (POTENTIAL).  
 FT DOMAIN 261 277 EXTRACELLULAR (POTENTIAL).  
 FT TRANSMEM 278 301 7 (POTENTIAL).  
 FT DOMAIN 302 352 CYTOPLASMIC (POTENTIAL).  
 FT DISULFID 101 178 BY SIMILARITY.



FT MOD\_RES 3 3 SULFATION (BY SIMILARITY).  
 FT MOD\_RES 10 10 SULFATION (BY SIMILARITY).  
 FT MOD\_RES 14 14 SULFATION (BY SIMILARITY).  
 FT MOD\_RES 15 15 SULFATION (BY SIMILARITY).  
 SQ SEQUENCE 352 AA; 40585 MW; 4366F142730F938F CRC64;

Query Match 45.8%; Score 365; DB 1; Length 352;  
 Best Local Similarity 33.6%; Pred. No. 4.1e-28;  
 Matches 95; Conservative 10; Mismatches 28; Indels 150; Gaps 7;

Qy 1 MDYQVSSPIYDINYYTSEPCQKINVKQIAA----- 30  
 ||||| |||:|||||||:|||||  
 Db 1 MDYQVSSPTYDIDYYTSEPCQKVNKQIAARLLPPLYSLVFIFGFVGNILVVLILINCKR 60  
 Qy 31 -----YKCGLC-----AAQWDFGNTMCQHQRVHGHHHSYKCG--- 64  
 | | ||||| | : | : |  
 Db 61 LKSMTDIYLLNLAISDLFFLLTVPFWAHYAAQWDFGNTMC--QLLTGLYFIGFFSGIFF 118  
 Qy 65 -----LCTRSQKEGLHYTC 78  
 : |||:|||||  
 Db 119 IILLTIDRYLAIVHAVFALKARTVTFGVTVSVITWVAVFASLPGIIFTRSQREGLHYTC 178  
 Qy 79 SSHFPYSQYQFWKNFQTLKI-----HQRVHGG----- 105  
 ||||| ||||| : : |  
 Db 179 SSHFPYSQYQFWKNFQTLKIVILGLVPLLVVICYSGLKTLRLCRNEKKRHRAVRLIF 238  
 Qy 106 -----GGSYKCGLC----QEFGFLNNCSSSNRLDGHQRV 135  
 | | ||||| : |  
 Db 239 TIMIVYFLFWAPYNIVLLNLTQFEFGFLNNCSSSNRLDQAMQV 281

# RESULT 7

CKR5\_PYGNE

ID CKR5\_PYGNE STANDARD; PRT; 352 AA.

AC 097882;

DT 30-MAY-2000 (Rel. 39, Created)

DT 30-MAY-2000 (Rel. 39, Last sequence update)

DT 28-FEB-2003 (Rel. 41, Last annotation update)

DE C-C chemokine receptor type 5 (C-C CKR-5) (CC-CKR-5) (CCR-5) (CCR5).

GN CCR5 OR CMKBR5.

OS Pygathrix nemaeus (Dove langur).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Primates; Catarrhini; Cercopithecidae; Colobinae;

OC Pygathrix.

OX NCBI\_TaxID=54133;

RN [1]

RP SEQUENCE FROM N.A.

RX MEDLINE=99416438; PubMed=10486970;

RA Zhang Y.-W., Ryder O.A., Zhang Y.-P.;

RT "Sequence evolution of the CCR5 chemokine receptor gene in primates.";

RL Mol. Biol. Evol. 16:1145-1154(1999).

CC -!- FUNCTION: Receptor for a C-C type chemokine. Binds to MIP-1-alpha,  
 CC MIP-1-beta and RANTES and subsequently transduces a signal by  
 CC increasing the intracellular calcium ions level. May play a role  
 CC in the control of granulocytic lineage proliferation or  
 CC differentiation.

CC -!- SUBCELLULAR LOCATION: Integral membrane protein.



Db 179 SSHFPYSQYQFWKNFQTLKIVILGLVLPLLLIMVICYSGILKTLRLCRNEKKRHRAVRLIF 238

Qy 106 -----GGSYKCGLC-----QEFFGLNNCSSSNRLDGHQRV 135  
| | ||||| :|

Db 239 TIMIVYFLEWAPYNIVLLLNTFQEFFGLNNCSSSNRLDQAMQV 281

RESULT 8

CKR5\_TRAFR

ID CKR5\_TRAFR STANDARD; PRT; 352 AA.

AC O97878;

DT 30-MAY-2000 (Rel. 39, Created)

DT 30-MAY-2000 (Rel. 39, Last sequence update)

DT 28-FEB-2003 (Rel. 41, Last annotation update)

DE C-C chemokine receptor type 5 (C-C CKR-5) (CC-CKR-5) (CCR-5) (CCR5).

GN CCR5 OR CMKBR5.

OS Trachypithecus francoisi (Francois' langur) (Indochinese langur).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Primates; Catarrhini; Cercopithecidae; Colobinae;

OC Trachypithecus.

OX NCBI\_TaxID=54180;

RN [1]

RP SEQUENCE FROM N.A.

RX MEDLINE=99416438; PubMed=10486970;

RA Zhang Y.-W., Ryder O.A., Zhang Y.-P.;

RT "Sequence evolution of the CCR5 chemokine receptor gene in primates.";

RL Mol. Biol. Evol. 16:1145-1154(1999).

CC -!- FUNCTION: Receptor for a C-C type chemokine. Binds to MIP-1-alpha,  
CC MIP-1-beta and RANTES and subsequently transduces a signal by  
CC increasing the intracellular calcium ions level. May play a role  
CC in the control of granulocytic lineage proliferation or  
CC differentiation.

CC -!- SUBCELLULAR LOCATION: Integral membrane protein.

CC -!- SIMILARITY: Belongs to family 1 of G-protein coupled receptors.

CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----

DR EMBL; AF075442; AAD19854.1; -.

DR InterPro; IPR000276; GPCR\_Rhodpsn.

DR Pfam; PF00001; 7tm\_1; 1.

DR PRINTS; PR00237; GPCRRHODOPSN.

DR PROSITE; PS00237; G\_PROTEIN\_RECEP\_F1\_1; 1.

DR PROSITE; PS50262; G\_PROTEIN\_RECEP\_F1\_2; 1.

KW G-protein coupled receptor; Transmembrane; Glycoprotein; Sulfation.

FT DOMAIN 1 30 EXTRACELLULAR (POTENTIAL).

FT TRANSMEM 31 58 1 (POTENTIAL).

FT DOMAIN 59 68 CYTOPLASMIC (POTENTIAL).

FT TRANSMEM 69 89 2 (POTENTIAL).

FT DOMAIN 90 102 EXTRACELLULAR (POTENTIAL).

FT TRANSMEM 103 124 3 (POTENTIAL).

FT DOMAIN 125 141 CYTOPLASMIC (POTENTIAL).

FT	TRANSMEM	142	166	4 (POTENTIAL).
FT	DOMAIN	167	198	EXTRACELLULAR (POTENTIAL).
FT	TRANSMEM	199	218	5 (POTENTIAL).
FT	DOMAIN	219	235	CYTOPLASMIC (POTENTIAL).
FT	TRANSMEM	236	260	6 (POTENTIAL).
FT	DOMAIN	261	277	EXTRACELLULAR (POTENTIAL).
FT	TRANSMEM	278	301	7 (POTENTIAL).
FT	DOMAIN	302	352	CYTOPLASMIC (POTENTIAL).
FT	DISULFID	101	178	BY SIMILARITY.
FT	MOD_RES	3	3	SULFATION (BY SIMILARITY).
FT	MOD_RES	10	10	SULFATION (BY SIMILARITY).
FT	MOD_RES	14	14	SULFATION (BY SIMILARITY).
FT	MOD_RES	15	15	SULFATION (BY SIMILARITY).
SQ	SEQUENCE	352 AA;	40496 MW;	4366F148C255938F CRC64;

Query Match 45.8%; Score 365; DB 1; Length 352;

Best Local Similarity 33.6%; Pred. No. 4.1e-28;

Matches 95; Conservative 10; Mismatches 28; Indels 150; Gaps 7;

Qy	1	MDYQVSSPIYDINYYTSEPCQKINVQIAA-----	30
		:       :	
Db	1	MDYQVSSPTYDIDYYTSEPCQKVNQIAARLLPPLYSLVFIFGFVGNILVVLILINCKR	60
Qy	31	-----YKCGLC-----AAAQWDFGNTMCQHQRVHGHHSYKCG---	64
		:   :	
Db	61	LKSMTDIYLLNLAISDLFFLLTVPFWAHYAAAQWDFGNTMC--QLLTGLYFIGFFSGIFF	118
Qy	65	-----LCTRSQKEGLHYTC	78
		:     :	
Db	119	IILLTIDRYLAIVHAVFALKARTVTFGVVTSVITWVAVFASLPGIIFTRSQREGLHYTC	178
Qy	79	SSHFPYSQYQFWKNFQTLKI-----HQRVHGG-----	105
		: :	
Db	179	SSHFPYSQYQFWKNFQTLKIVILGLVLPLLVMVICYSGILKTLRLCRNEKKRHRAVRLIF	238
Qy	106	-----GGSYKCGLC----QEFFGLNNCSSSNRLDGHQRV	135
		:	
Db	239	TIMIVYFLFWAPYNIVLLLNTFQEFFGLNNCSSSNRLDQAMQV	281

# RESULT 9

CKR5\_TRAPH

ID CKR5\_TRAPH STANDARD; PRT; 352 AA.

AC 097879;

DT 30-MAY-2000 (Rel. 39, Created)

DT 30-MAY-2000 (Rel. 39, Last sequence update)

DT 28-FEB-2003 (Rel. 41, Last annotation update)

DE C-C chemokine receptor type 5 (C-C CKR-5) (CC-CKR-5) (CCR-5) (CCR5).

GN CCR5 OR CMKBR5.

OS Trachypithecus phayrei (Phayre's leaf monkey).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Primates; Catarrhini; Cercopithecidae; Colobinae;

OC Trachypithecus.

OX NCBI\_TaxID=61618;

RN [1]

RP SEQUENCE FROM N.A.

RX MEDLINE=99416438; PubMed=10486970;

RA Zhang Y.-W., Ryder O.A., Zhang Y.-P.;  
 RT "Sequence evolution of the CCR5 chemokine receptor gene in primates.";  
 RL Mol. Biol. Evol. 16:1145-1154(1999).  
 CC -!- FUNCTION: Receptor for a C-C type chemokine. Binds to MIP-1-alpha,  
 CC MIP-1-beta and RANTES and subsequently transduces a signal by  
 CC increasing the intracellular calcium ions level. May play a role  
 CC in the control of granulocytic lineage proliferation or  
 CC differentiation.  
 CC -!- SUBCELLULAR LOCATION: Integral membrane protein.  
 CC -!- SIMILARITY: Belongs to family 1 of G-protein coupled receptors.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 DR EMBL; AF075443; AAD19855.1; -.  
 DR InterPro; IPR000276; GPCR\_Rhodpsn.  
 DR Pfam; PF00001; 7tm\_1; 1.  
 DR PRINTS; PR00237; GPCRRHODOPSN.  
 DR PROSITE; PS00237; G\_PROTEIN\_RECEP\_F1\_1; 1.  
 DR PROSITE; PS50262; G\_PROTEIN\_RECEP\_F1\_2; 1.  
 KW G-protein coupled receptor; Transmembrane; Glycoprotein; Sulfation.  
 FT DOMAIN 1 30 EXTRACELLULAR (POTENTIAL).  
 FT TRANSMEM 31 58 1 (POTENTIAL).  
 FT DOMAIN 59 68 CYTOPLASMIC (POTENTIAL).  
 FT TRANSMEM 69 89 2 (POTENTIAL).  
 FT DOMAIN 90 102 EXTRACELLULAR (POTENTIAL).  
 FT TRANSMEM 103 124 3 (POTENTIAL).  
 FT DOMAIN 125 141 CYTOPLASMIC (POTENTIAL).  
 FT TRANSMEM 142 166 4 (POTENTIAL).  
 FT DOMAIN 167 198 EXTRACELLULAR (POTENTIAL).  
 FT TRANSMEM 199 218 5 (POTENTIAL).  
 FT DOMAIN 219 235 CYTOPLASMIC (POTENTIAL).  
 FT TRANSMEM 236 260 6 (POTENTIAL).  
 FT DOMAIN 261 277 EXTRACELLULAR (POTENTIAL).  
 FT TRANSMEM 278 301 7 (POTENTIAL).  
 FT DOMAIN 302 352 CYTOPLASMIC (POTENTIAL).  
 FT DISULFID 101 178 BY SIMILARITY.  
 FT MOD\_RES 3 3 SULFATION (BY SIMILARITY).  
 FT MOD\_RES 10 10 SULFATION (BY SIMILARITY).  
 FT MOD\_RES 14 14 SULFATION (BY SIMILARITY).  
 FT MOD\_RES 15 15 SULFATION (BY SIMILARITY).  
 SQ SEQUENCE 352 AA; 40509 MW; 4366F148D3A5938F CRC64;

Query Match 45.8%; Score 365; DB 1; Length 352;  
 Best Local Similarity 33.6%; Pred. No. 4.1e-28;  
 Matches 95; Conservative 10; Mismatches 28; Indels 150; Gaps 7;

Qy 1 MDYQVSSPIYDINYTTSEPCQKINVKQIAA----- 30  
 ||||| |||:|||||:|||||  
 Db 1 MDYQVSSPTYDIDYYTSEPCQKVNKQIAARLLPLYSLVFIFGFVGNILVVLILINCKR 60  
 Qy 31 -----YKGLC-----AAQWDFGNTMCQHQRVHGHHSYKCG--- 64

```

Db          61  LKSMTDIYLLNLAIISDLFFLLTVPFWAHYAAAQWDFGNTMC--QLLTGLYFIGFFSGIFF 118
Qy          65  -----LCTRSQKEGLHYTC 78
              : ||||:|||||
Db          119 IILLTIDRYLAIVHAVFALKARTVTFGVVTSVITWVVAVFASLPGIIFTRSQREGLHYTC 178
Qy          79  SSHFPYSQYQFWKNFQTLKI-----HQRVHGG----- 105
              |||||:|
Db          179 SSHFPYSQYQFWKNFQTLKIVILGLVLP LLVMVICYSGILKTLLRCRNEKKRHRAVRLIF 238
Qy          106 -----GGSYKCGLC----QEFFGLNNCSSSNRLDGHQRV 135
              | | |||||:|
Db          239 TIMIVYFLFWAPYNIVLLLNTEQEFFGLNNCSSSNRLDQAMQV 281

```

DR PRINTS; PR00237; GPCRRHODOPSN.  
 DR PROSITE; PS00237; G\_PROTEIN\_RECEP\_F1\_1; 1.  
 DR PROSITE; PS50262; G\_PROTEIN\_RECEP\_F1\_2; 1.  
 KW G-protein coupled receptor; Transmembrane; Glycoprotein; Sulfation.  
 FT DOMAIN 1 30 EXTRACELLULAR (POTENTIAL).  
 FT TRANSMEM 31 58 1 (POTENTIAL).  
 FT DOMAIN 59 68 CYTOPLASMIC (POTENTIAL).  
 FT TRANSMEM 69 89 2 (POTENTIAL).  
 FT DOMAIN 90 102 EXTRACELLULAR (POTENTIAL).  
 FT TRANSMEM 103 124 3 (POTENTIAL).  
 FT DOMAIN 125 141 CYTOPLASMIC (POTENTIAL).  
 FT TRANSMEM 142 166 4 (POTENTIAL).  
 FT DOMAIN 167 198 EXTRACELLULAR (POTENTIAL).  
 FT TRANSMEM 199 218 5 (POTENTIAL).  
 FT DOMAIN 219 235 CYTOPLASMIC (POTENTIAL).  
 FT TRANSMEM 236 260 6 (POTENTIAL).  
 FT DOMAIN 261 277 EXTRACELLULAR (POTENTIAL).  
 FT TRANSMEM 278 301 7 (POTENTIAL).  
 FT DOMAIN 302 352 CYTOPLASMIC (POTENTIAL).  
 FT DISULFID 101 178 BY SIMILARITY.  
 FT MOD\_RES 3 3 SULFATION (BY SIMILARITY).  
 FT MOD\_RES 10 10 SULFATION (BY SIMILARITY).  
 FT MOD\_RES 14 14 SULFATION (BY SIMILARITY).  
 FT MOD\_RES 15 15 SULFATION (BY SIMILARITY).  
 SQ SEQUENCE 352 AA; 40515 MW; D0E6FCB9FE5EAC84 CRC64;

Query Match 45.7%; Score 364; DB 1; Length 352;  
 Best Local Similarity 33.9%; Pred. No. 5.1e-28;  
 Matches 96; Conservative 8; Mismatches 29; Indels 150; Gaps 7;

Qy 1 MDYQVSSPIYDINYYTSEPCQKINVKQIAA----- 30  
 ||||| : |||||  
 Db 1 MDYQVSSPTYDIDYYTSEPCQKTNVKQIAARLLPPLYSLVFIFGFVGNMLVILILINCKR 60  
 Qy 31 -----YKCGLC-----AAAQWDFGNTMCQHQRVHGHHSYKCG--- 64  
 | | ||||| : | : |  
 Db 61 LKSMTDIYLLNLAISDLFFLLTPFEWAHAAAQWDFGNTMC--QLLTGLYFIGFFSGIFF 118  
 Qy 65 -----LCTRSQKEGLHYTC 78  
 : |||||  
 Db 119 IILLTIDRYLAIVHAVFALKARTVTFGVVTSVITWVAVFASLPGIIFTRSQKEGLHYTC 178  
 Qy 79 SSHFPYSQYQFWKNFQTLKI-----HQRVHGG----- 105  
 ||||| : |  
 Db 179 SSHFPYSQYQFWKNFQTLKIVILGLVLP LLVMVICYS GILKTL LRCRNEKKRHRVRLIF 238  
 Qy 106 -----GGSYKCGLC----QEFFGLNNC SSSNRLDGHQRV 135  
 | | ||||| : |  
 Db 239 TIMIVYFLFWAPYNIVLLNLT FQEFFGLNNC SSSNRLDQAMQV 281

# RESULT 11

CKR5\_MACMU

ID CKR5\_MACMU STANDARD; PRT; 352 AA.

AC P79436; 002746;

DT 01-NOV-1997 (Rel. 35, Created)

DT 01-NOV-1997 (Rel. 35, Last sequence update)

DT 28-FEB-2003 (Rel. 41, Last annotation update)  
 DE C-C chemokine receptor type 5 (C-C CKR-5) (CC-CKR-5) (CCR-5) (CCR5).  
 GN CCR5 OR CMKBR5.  
 OS *Macaca mulatta* (Rhesus macaque),  
 OS *Macaca fascicularis* (Crab eating macaque) (*Cynomolgus* monkey), and  
 OS *Macaca nemestrina* (Pig-tailed macaque).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Cercopithecidae;  
 OC Cercopithecinae; *Macaca*.  
 OX NCBI\_TaxID=9544, 9541, 9545;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC SPECIES=*M.mulatta*;  
 RX MEDLINE=97184592; PubMed=9032394;  
 RA Marcon L., Choe H., Martin K.A., Farzan M., Ponath P.D., Wu L.,  
 RA Newman W., Gerard N., Gerard C., Sodroski J.;  
 RT "Utilization of C-C chemokine receptor 5 by the envelope  
 RT glycoproteins of a pathogenic simian immunodeficiency virus,  
 RT SIVmac239.";  
 RL J. Virol. 71:2522-2527(1997).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RC SPECIES=*M.mulatta*; STRAIN=Indian macaque;  
 RX MEDLINE=97213934; PubMed=9060623;  
 RA Chen Z., Zhou P., Ho D.D., Landau N.R., Marx P.A.;  
 RT "Genetically divergent strains of simian immunodeficiency virus use  
 RT CCR5 as a coreceptor for entry.";  
 RL J. Virol. 71:2705-2714(1997).  
 RN [3]  
 RP SEQUENCE FROM N.A.  
 RC SPECIES=*M.mulatta*;  
 RX MEDLINE=21354176; PubMed=11461684;  
 RA Margulies B.J., Hauer D.A., Clements J.E.;  
 RT "Identification and comparison of eleven rhesus macaque chemokine  
 RT receptors.";  
 RL AIDS Res. Hum. Retroviruses 17:981-986(2001).  
 RN [4]  
 RP SEQUENCE FROM N.A.  
 RC SPECIES=*M.mulatta*, *M.fascicularis*, and *M.nemestrina*;  
 RX MEDLINE=97268687; PubMed=9108095;  
 RA Edinger A.L., Amedee A., Miller K., Doranz B.J., Endres M.,  
 RA Sharron M., Samson M., Lu Z.H., Clements J.E., Murphey-Corb M.,  
 RA Peiper S.C., Parmentier M., Broder C.C., Doms R.W.;  
 RT "Differential utilization of CCR5 by macrophage and T cell tropic  
 RT simian immunodeficiency virus strains.";  
 RL Proc. Natl. Acad. Sci. U.S.A. 94:4005-4010(1997).  
 CC -!- FUNCTION: Receptor for a C-C type chemokine. Binds to MIP-1-alpha,  
 CC MIP-1-beta and RANTES and subsequently transduces a signal by  
 CC increasing the intracellular calcium ions level. May play a role  
 CC in the control of granulocytic lineage proliferation or  
 CC differentiation.  
 CC -!- SUBCELLULAR LOCATION: Integral membrane protein.  
 CC -!- SIMILARITY: Belongs to family 1 of G-protein coupled receptors.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its



CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

-----  
 DR EMBL; U77672; AAC51109.1; -.  
 DR EMBL; U73739; AAC51158.1; -.  
 DR EMBL; U96762; AAC34132.1; -.  
 DR EMBL; AF005660; AAB62554.1; -.  
 DR EMBL; AF005661; AAB62555.1; -.  
 DR EMBL; AF005662; AAB62556.1; -.  
 DR InterPro; IPR000276; GPCR\_Rhodpsn.  
 DR Pfam; PF00001; 7tm\_1; 1.  
 DR PRINTS; PR00237; GPCRRHODOPSN.  
 DR PROSITE; PS00237; G\_PROTEIN\_RECEP\_F1\_1; 1.  
 DR PROSITE; PS50262; G\_PROTEIN\_RECEP\_F1\_2; 1.  
 KW G-protein coupled receptor; Transmembrane; Glycoprotein; Sulfation.  
 FT DOMAIN 1 30 EXTRACELLULAR (POTENTIAL).  
 FT TRANSMEM 31 58 1 (POTENTIAL).  
 FT DOMAIN 59 68 CYTOPLASMIC (POTENTIAL).  
 FT TRANSMEM 69 89 2 (POTENTIAL).  
 FT DOMAIN 90 102 EXTRACELLULAR (POTENTIAL).  
 FT TRANSMEM 103 124 3 (POTENTIAL).  
 FT DOMAIN 125 141 CYTOPLASMIC (POTENTIAL).  
 FT TRANSMEM 142 166 4 (POTENTIAL).  
 FT DOMAIN 167 198 EXTRACELLULAR (POTENTIAL).  
 FT TRANSMEM 199 218 5 (POTENTIAL).  
 FT DOMAIN 219 235 CYTOPLASMIC (POTENTIAL).  
 FT TRANSMEM 236 260 6 (POTENTIAL).  
 FT DOMAIN 261 277 EXTRACELLULAR (POTENTIAL).  
 FT TRANSMEM 278 301 7 (POTENTIAL).  
 FT DOMAIN 302 352 CYTOPLASMIC (POTENTIAL).  
 FT DISULFID 101 178 BY SIMILARITY.  
 FT MOD\_RES 3 3 SULFATION (BY SIMILARITY).  
 FT MOD\_RES 10 10 SULFATION (BY SIMILARITY).  
 FT MOD\_RES 14 14 SULFATION (BY SIMILARITY).  
 FT MOD\_RES 15 15 SULFATION (BY SIMILARITY).  
 FT CARBOHYD 268 268 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CONFLICT 241 241 M -> I (IN REF. 3).  
 FT CONFLICT 292 292 I -> M (IN REF. 3).  
 SQ SEQUENCE 352 AA; 40507 MW; 58B96C85909FACB2 CRC64;

Query Match 45.5%; Score 363; DB 1; Length 352;  
 Best Local Similarity 33.6%; Pred. No. 6.4e-28;  
 Matches 95; Conservative 10; Mismatches 28; Indels 150; Gaps 7;

QY 1 MDYQVSSPIYDINYYTSEPCQKINVKQIAA----- 30  
 ||||| |||:|||||||  
 Db 1 MDYQVSSPTYDIDYYTSEPCQKINVKQIAARLLPPLYSLVFIFGFVGNILVVLILINCKR 60  
 QY 31 -----YKGLC-----AAAQWDFGNTMCQHQRVHGHHSYKCG--- 64  
 | | ||||| | : | : |  
 Db 61 LKSMTDIYLLNLAISDLLFLLTVPFWAHYAAAQWDFGNTMC--QLLTGLYFIGFFSGIFF 118  
 QY 65 -----LCTRSQKEGLHYTC 78  
 : |||:|||||  
 Db 119 IILLTIDRYLAIVHAVFALKARTVTFGVVTSVITWVAVFASLPGIIFTRSQREGHYTC 178

```

QY      79 SSHFPYSQYQFWKNFQTLKI-----HQRVHGG----- 105
      |||||:
Db      179 SSHFPYSQYQFWKNFQTLKMVILGLVPLLVVICYSGLKTLRLCRNEKKRHRVRLIF 238

QY      106 -----GGSYKCGLC----QEFFGLNNCSSSNRLDGHQRV 135
      | | |||||:|
Db      239 TIMIVYFLFWAPYNIVLLNTFQEFFGLNNCSSSNRLDQAMQV 281

```

# RESULT 12

## CKR5\_HYLML

```

ID      CKR5_HYLML      STANDARD;      PRT;      352 AA.
AC      Q95NC0;
DT      15-MAR-2004 (Rel. 43, Created)
DT      15-MAR-2004 (Rel. 43, Last sequence update)
DT      15-MAR-2004 (Rel. 43, Last annotation update)
DE      C-C chemokine receptor type 5 (C-C CKR-5) (CC-CKR-5) (CCR-5) (CCR5).
GN      CCR5 OR CMKBR5.
OS      Hylobates moloch (Silvery gibbon).
OC      Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC      Mammalia; Eutheria; Primates; Catarrhini; Hylobatidae; Hylobates.
OX      NCBI_TaxID=81572;
RN      [1]
RP      SEQUENCE FROM N.A.
RX      MEDLINE=99416438; PubMed=10486970;
RA      Zhang Y.-W., Ryder O.A., Zhang Y.-P.;
RT      "Sequence evolution of the CCR5 chemokine receptor gene in primates.";
RL      Mol. Biol. Evol. 16:1145-1154(1999).
CC      -!- FUNCTION: Receptor for a C-C type chemokine. Binds to MIP-1-alpha,
CC      MIP-1-beta and RANTES and subsequently transduces a signal by
CC      increasing the intracellular calcium ions level. May play a role
CC      in the control of granulocytic lineage proliferation or
CC      differentiation.
CC      -!- SUBCELLULAR LOCATION: Integral membrane protein.
CC      -!- SIMILARITY: Belongs to family 1 of G-protein coupled receptors.
CC      -----
CC      This SWISS-PROT entry is copyright. It is produced through a collaboration
CC      between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC      the European Bioinformatics Institute. There are no restrictions on its
CC      use by non-profit institutions as long as its content is in no way
CC      modified and this statement is not removed. Usage by and for commercial
CC      entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC      or send an email to license@isb-sib.ch).
CC      -----
DR      EMBL; AF177899; AAK43382.1; -.
DR      InterPro; IPR000276; GPCR_Rhodpsn.
DR      Pfam; PF00001; 7tm_1; 1.
DR      PRINTS; PR00237; GPCRRHODOPSN.
DR      PROSITE; PS00237; G_PROTEIN_RECEP_F1_1; 1.
DR      PROSITE; PS50262; G_PROTEIN_RECEP_F1_2; 1.
KW      G-protein coupled receptor; Transmembrane; Glycoprotein; Sulfation.
FT      DOMAIN      1      30      EXTRACELLULAR (POTENTIAL).
FT      TRANSMEM      31      58      1 (POTENTIAL).
FT      DOMAIN      59      68      CYTOPLASMIC (POTENTIAL).
FT      TRANSMEM      69      89      2 (POTENTIAL).
FT      DOMAIN      90      102      EXTRACELLULAR (POTENTIAL).

```

FT	TRANSMEM	103	124	3 (POTENTIAL).
FT	DOMAIN	125	141	CYTOPLASMIC (POTENTIAL).
FT	TRANSMEM	142	166	4 (POTENTIAL).
FT	DOMAIN	167	198	EXTRACELLULAR (POTENTIAL).
FT	TRANSMEM	199	218	5 (POTENTIAL).
FT	DOMAIN	219	235	CYTOPLASMIC (POTENTIAL).
FT	TRANSMEM	236	260	6 (POTENTIAL).
FT	DOMAIN	261	277	EXTRACELLULAR (POTENTIAL).
FT	TRANSMEM	278	301	7 (POTENTIAL).
FT	DOMAIN	302	352	CYTOPLASMIC (POTENTIAL).
FT	DISULFID	101	178	BY SIMILARITY.
FT	MOD_RES	3	3	SULFATION (BY SIMILARITY).
FT	MOD_RES	10	10	SULFATION (BY SIMILARITY).
FT	MOD_RES	14	14	SULFATION (BY SIMILARITY).
SQ	SEQUENCE	352 AA;	40436 MW;	9623CA98340CF274 CRC64;

Query Match 45.4%; Score 362; DB 1; Length 352;

Best Local Similarity 33.9%; Pred. No. 8e-28;

Matches 96; Conservative 8; Mismatches 29; Indels 150; Gaps 7;

Qy	1	MDYQVSSPIYDINYYTSEPCQKINVKQIAA-----	30
		:	
Db	1	MDYQVSSPTYDIDYYTSGPCQKINVKQIAARLLPPLYSLVFIFGFVGNMLVILILINCKR	60
Qy	31	-----YKCGLC-----AAAQWDFGNTMCQHQRVHGHHSYKCG---	64
		:   :	
Db	61	LKSMTDIYLLNLAISDLFFLLTVPFWAHYAAAQWDFGNTMC--QLLTGLYFIGFFSGIFF	118
Qy	65	-----LCTRSQKEGLHYTC	78
		:	
Db	119	IILLTIDRYLAIVHAVEFALKARTVTFGVTVSVITWVAVFASLPGLIIFTRSQKEGLHYTC	178
Qy	79	SSHFPYSQYQFWKNFQTLKI-----HQRVHGG-----	105
		:	
Db	179	SSHFPYSQYQFWKNFQTLKIVILGLVLPPLVMVICYSGLKTLRLCRNEKKRHRVRLIF	238
Qy	106	-----GGSYKCGLC---QEFGNLCSSSNRLDGHQRV	135
		:	
Db	239	TIMIVYFLEWAPYNIVLLLNTFQEFGNLCSSSNRLDQAMQV	281

# RESULT 13

CKR5\_CERP

ID CKR5\_CERP STANDARD; PRT; 352 AA.

AC Q9TV42;

DT 15-MAR-2004 (Rel. 43, Created)

DT 15-MAR-2004 (Rel. 43, Last sequence update)

DT 15-MAR-2004 (Rel. 43, Last annotation update)

DE C-C chemokine receptor type 5 (C-C CKR-5) (CC-CKR-5) (CCR-5) (CCR5).

GN CCR5 OR CMKBR5.

OS Cercopithecus pygerythrus (Vervet monkey).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Primates; Catarrhini; Cercopithecidae;

OC Cercopithecinae; Cercopithecus.

OX NCBI\_TaxID=60710;

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN=100;  
 RX MEDLINE=99335215; PubMed=10408730;  
 RA Muller-Trutwin M.C., Corbet S., Hansen J., Georges-Courbot M.-C.,  
 RA Diop O., Rigoulet J., Barre-Sinoussi F., Fomsgaard A.;  
 RT "Mutations in CCR5-coding sequences are not associated with SIV  
 RT carrier status in African nonhuman primates.";  
 RL AIDS Res. Hum. Retroviruses 15:931-939(1999).  
 CC -!- FUNCTION: Receptor for a C-C type chemokine. Binds to MIP-1-alpha,  
 CC MIP-1-beta and RANTES and subsequently transduces a signal by  
 CC increasing the intracellular calcium ions level. May play a role  
 CC in the control of granulocytic lineage proliferation or  
 CC differentiation.  
 CC -!- SUBCELLULAR LOCATION: Integral membrane protein.  
 CC -!- SIMILARITY: Belongs to family 1 of G-protein coupled receptors.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 DR EMBL; AF035222; AAD44015.1; -.  
 DR InterPro; IPR000276; GPCR\_Rhodpsn.  
 DR Pfam; PF00001; 7tm\_1; 1.  
 DR PRINTS; PR00237; GPCRRHODOPSN.  
 DR PROSITE; PS00237; G\_PROTEIN\_RECEP\_F1\_1; 1.  
 DR PROSITE; PS50262; G\_PROTEIN\_RECEP\_F1\_2; 1.  
 KW G-protein coupled receptor; Transmembrane; Glycoprotein; Sulfation.  
 FT DOMAIN 1 30 EXTRACELLULAR (POTENTIAL).  
 FT TRANSMEM 31 58 1 (POTENTIAL).  
 FT DOMAIN 59 68 CYTOPLASMIC (POTENTIAL).  
 FT TRANSMEM 69 89 2 (POTENTIAL).  
 FT DOMAIN 90 102 EXTRACELLULAR (POTENTIAL).  
 FT TRANSMEM 103 124 3 (POTENTIAL).  
 FT DOMAIN 125 141 CYTOPLASMIC (POTENTIAL).  
 FT TRANSMEM 142 166 4 (POTENTIAL).  
 FT DOMAIN 167 198 EXTRACELLULAR (POTENTIAL).  
 FT TRANSMEM 199 218 5 (POTENTIAL).  
 FT DOMAIN 219 235 CYTOPLASMIC (POTENTIAL).  
 FT TRANSMEM 236 260 6 (POTENTIAL).  
 FT DOMAIN 261 277 EXTRACELLULAR (POTENTIAL).  
 FT TRANSMEM 278 301 7 (POTENTIAL).  
 FT DOMAIN 302 352 CYTOPLASMIC (POTENTIAL).  
 FT DISULFID 101 178 BY SIMILARITY.  
 FT MOD\_RES 3 3 SULFATION (BY SIMILARITY).  
 FT MOD\_RES 10 10 SULFATION (BY SIMILARITY).  
 FT MOD\_RES 15 15 SULFATION (BY SIMILARITY).  
 SQ SEQUENCE 352 AA; 40558 MW; EF17D67CBCCC3DB0 CRC64;

Query Match 45.3%; Score 361; DB 1; Length 352;

Best Local Similarity 33.6%; Pred. No. 9.9e-28;

Matches 95; Conservative 9; Mismatches 29; Indels 150; Gaps 7;

QY 1 MDYQVSSPIYDINYTTSEPCQKINVKQIAA----- 30  
 |||:|||||

```

Db          1 MDYQVSSPTYDIDYYTSEPCQKINVKQIAARLLPPLYSLVFIFGFVGNILVVLILINCKR 60
Qy          31 -----YKCGLC-----AAAQWDFGNTMCQHQRVHGHHHSYKCG--- 64
              |   |               ||||| | | | | | | | | | | : | : | : |
Db          61 LKSMTDIYLLNLAIISDLLFLLTVPFWAHYAAAQWDFGNTMC--QLLTGLYFIGFFSGIFF 118
Qy          65 -----LCTRSQKEGLHYTC 78
                      : ||||:|||| |
Db          119 IILLTIDRYLAIVHAVFALKARTVTFGVVTSVITWVVAVFASLPRIIFTRSQREGLHYAC 178
Qy          79 SSHFPYSQYQFWKNFQTLKI-----HQRVHGG----- 105
              ||||| | | | | | | | | | | | | | | | | | | | : : |
Db          179 SSHFPYSQYQFWKNFQTLKIVILGLVLP LLVMVICYS GILKTL LRCRNEKKRHRVRLIF 238
Qy          106 -----GGSYKCGLC----QEFFGLNNC SSSNRLDGHQRV 135
              |   |               ||||| | | | | | | | | | | : |
Db          239 TIMIVYFLFWAPYNIVLLNTFQEFFGLNNC SSSNRLDQAMQV 281

```

#### RESULT 14

##### CKR5\_CERTO

```

ID   CKR5_CERTO      STANDARD;          PRT;   352 AA.
AC   O62743; O62744; O62745; O62746;
DT   15-DEC-1998 (Rel. 37, Created)
DT   15-DEC-1998 (Rel. 37, Last sequence update)
DT   16-OCT-2001 (Rel. 40, Last annotation update)
DE   C-C chemokine receptor type 5 (C-C CKR-5) (CC-CKR-5) (CCR-5) (CCR5).
GN   CCR5 OR CMKBR5.
OS   Cercopithecus torquatus atys (Red-crowned mangabey) (Sooty mangabey).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Cercopithecidae;
OC   Cercopithecinae; Cercopithecus.
OX   NCBI_TaxID=9531;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=Isolate 079, 085, 087, and 089;
RX   MEDLINE=98321155; PubMed=9656999;
RA   Chen Z., Gettie A., Ho D.D., Marx P.A.;
RT   "Primary SIVsm isolates use the CCR5 coreceptor from sooty mangabeys
RT   naturally infected in west Africa: a comparison of coreceptor usage
RT   of primary SIVsm, HIV-2, and SIVmac.";
RL   Virology 246:113-124(1998).
CC   -!- FUNCTION: Receptor for a C-C type chemokine. Binds to MIP-1-alpha,
CC   MIP-1-beta and RANTES and subsequently transduces a signal by
CC   increasing the intracellular calcium ions level. May play a role
CC   in the control of granulocytic lineage proliferation or
CC   differentiation.
CC   -!- SUBCELLULAR LOCATION: Integral membrane protein.
CC   -!- SIMILARITY: Belongs to family 1 of G-protein coupled receptors.
CC   -----
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).

```

```

CC -----
DR EMBL; AF051902; AAC39830.1; -.
DR EMBL; AF051903; AAC39831.1; -.
DR EMBL; AF051904; AAC39832.1; -.
DR EMBL; AF051905; AAC39833.1; -.
DR InterPro; IPR000276; GPCR_Rhodpsn.
DR Pfam; PF00001; 7tm_1; 1.
DR PRINTS; PR00237; GPCRRHODOPSN.
DR PROSITE; PS00237; G_PROTEIN_RECEP_F1_1; 1.
DR PROSITE; PS50262; G_PROTEIN_RECEP_F1_2; 1.
KW G-protein coupled receptor; Transmembrane; Glycoprotein; Sulfation.
FT DOMAIN 1 30 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 31 58 1 (POTENTIAL).
FT DOMAIN 59 68 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 69 89 2 (POTENTIAL).
FT DOMAIN 90 102 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 103 124 3 (POTENTIAL).
FT DOMAIN 125 141 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 142 166 4 (POTENTIAL).
FT DOMAIN 167 198 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 199 218 5 (POTENTIAL).
FT DOMAIN 219 235 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 236 260 6 (POTENTIAL).
FT DOMAIN 261 277 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 278 301 7 (POTENTIAL).
FT DOMAIN 302 352 CYTOPLASMIC (POTENTIAL).
FT DISULFID 101 178 BY SIMILARITY.
FT MOD_RES 3 3 SULFATION (BY SIMILARITY).
FT MOD_RES 10 10 SULFATION (BY SIMILARITY).
FT MOD_RES 14 14 SULFATION (BY SIMILARITY).
FT MOD_RES 15 15 SULFATION (BY SIMILARITY).
FT VARIANT 2 2 D -> E (IN ISOLATE 087).
FT VARIANT 3 3 Y -> D (IN ISOLATE 079).
FT VARIANT 25 25 V -> G (IN ISOLATE 087).
FT VARIANT 100 100 M -> K (IN ISOLATE 079).
FT VARIANT 107 107 L -> V (IN ISOLATE 089).
FT VARIANT 134 134 V -> G (IN ISOLATE 079).
FT VARIANT 146 146 V -> L (IN ISOLATES 085 AND 089).
FT VARIANT 340 340 T -> I (IN ISOLATE 079).
SQ SEQUENCE 352 AA; 40489 MW; 20A196E2D47E49CA CRC64;

```

```

Query Match 45.3%; Score 361; DB 1; Length 352;
Best Local Similarity 33.6%; Pred. No. 9.9e-28;
Matches 95; Conservative 9; Mismatches 29; Indels 150; Gaps 7;

```

```

Qy 1 MDYQVSSPIYDINYYTSEPCQKINVKQIAA----- 30
    ||||| |||:|||||
Db 1 MDYQVSSPTYDIDYYTSEPCQKINVKQIAARLLPPLYSLVFIFGFVGNILVVLILINCKR 60

Qy 31 -----YKCGLC-----AAAQWDFGNTMCQHQRVHGHHSYKCG--- 64
    | | | ||||| | : | : |
Db 61 LKSMTDIYLLNLAISDLLFLLTVPFWAHYAAAQWDFGNTMC--QLLTGLYFIGFFSGIFF 118

Qy 65 -----LCTRSQKEGLHYTC 78
    : ||||:|||||
Db 119 IILLTIDRYLAIVHAVFALKARTVTFGVVTSVITWVAVFASLPGIIFTRSQKEGLHYTC 178

```



FT	DOMAIN	125	141	CYTOPLASMIC (POTENTIAL).
FT	TRANSMEM	142	166	4 (POTENTIAL).
FT	DOMAIN	167	198	EXTRACELLULAR (POTENTIAL).
FT	TRANSMEM	199	218	5 (POTENTIAL).
FT	DOMAIN	219	235	CYTOPLASMIC (POTENTIAL).
FT	TRANSMEM	236	260	6 (POTENTIAL).
FT	DOMAIN	261	277	EXTRACELLULAR (POTENTIAL).
FT	TRANSMEM	278	301	7 (POTENTIAL).
FT	DOMAIN	302	352	CYTOPLASMIC (POTENTIAL).
FT	DISULFID	101	178	BY SIMILARITY.
FT	MOD_RES	3	3	SULEFATION (BY SIMILARITY).
FT	MOD_RES	10	10	SULEFATION (BY SIMILARITY).
FT	MOD_RES	14	14	SULEFATION (BY SIMILARITY).
SQ	SEQUENCE	352 AA;	40445 MW;	4F8E4F344CEB7C91 CRC64;

Query Match 45.0%; Score 359; DB 1; Length 352;

Best Local Similarity 33.9%; Pred. No. 1.6e-27;

Matches 96; Conservative 8; Mismatches 29; Indels 150; Gaps 7;

Qy	1	MDYQVSSPIYDINYTSEPCQKINVQIAA-----	30
		:	
Db	1	MDYQVSSPTYDIDYDTSEPCQKINVQIAARLLPPLYSLVFIFGFVGNMLVILVLINCKR	60
Qy	31	-----YKCGLC-----AAAQWDFGNTMCQHQRVHGHHHSYKCG---	64
		:   :	
Db	61	LKSMTDIYLLNLAISDLFFLLTVPFWAHYAAAQWDFGNTMC--QLLTGLYFIGFFSGIFF	118
Qy	65	-----LCTRSQKEGLHYTC	78
		:	
Db	119	IILLTIDRYLAIVHAVFALKARTVTFGVVTSVITWVAVFASLPGIIFTRSQKEGLHYTC	178
Qy	79	SSHFPYSQYQFWKNFQTLKI-----HQRVHGG-----	105
		:	
Db	179	SSHFPYSQYQFWKNFQTLKIVILGLVLP LLVMVICYSGILKTLLRCRNEKKRHRVRLIF	238
Qy	106	-----GGSYKCGLC----QEFFGLNNCSSSNRLDGHQRV	135
		:	
Db	239	TIMIVYFLEWAPYNIVLLLNTFQEFFGLNNCSSSNRLDQAMQV	281

Search completed: September 17, 2004, 16:45:42

Job time : 15.0339 secs